# **DECLARATION**

I, Akiko KOSBMURA, of HIRAKI & ASSOCIATES, do solemnly and sincerely declare as follows:

- 1. That I am well acquainted with the English and Japanese languages and am competent to translate from Japanese into English.
- 2. That I have executed, with the best of my ability, a true and correct translation into English of Japanese Patent Application No. 148242/2003 filed on May 26, 2003, a copy of which I attach herewith.

This 16th day of July, 2010

A. Kaona Akiko KOSEMURA [Title of Document] DESCRIPTION

[Title Of Invention] A NUCLEIC ACID CONSTRUCT CONTAINING A NUCLEIC ACID DERIVED FROM THE GENOME OF HEPATITIS C VIRUS (HCV) OF GENOTYPE 2a, AND A CELL HAVING SUCH NUCLEIC ACID CONSTRUCT INTRODUCED THEREIN

[CLAIMS]

[Claim 1]

A replicon RNA, comprising a nucleotide sequence containing at least the 5' untranslated region, the sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein and the 3' untranslated region on the genomic RNA of hepatitis C virus of genotype 2a.

[Claim 2]

The replicon RNA of claim 1, further containing at least one selection marker gene or a reporter gene, and at least one IRES sequence.

[Claim 3]

A replicon RNA, comprising a nucleotide sequence containing the 5' untranslated region comprising the nucleotide sequence represented by SEQ ID NO: 9 or 10; at least one selection marker gene or a reporter gene; an IRES sequence; the sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein on the genomic RNA of hepatitis C virus of genotype 2a; and the 3' untranslated region comprising the nucleotide sequence represented by SEQ ID NO: 11 or 12.

[Claim 4]

The replicon RNA of any one of claims 1 to 3, wherein the genomic RNA of hepatitis C virus of genotype 2a is an RNA comprising the nucleotide sequence represented by SEQ ID NO: 3 or 5.

[Claim 5]

A replicon RNA, comprising the following RNA (a) or (b):

(a) an RNA comprising the nucleotide sequence represented by SEQ ID NO: 1 or

2; and

(b) an RNA comprising a nucleotide sequence derived from the nucleotide sequence represented by SEQ ID NO: 1 or 2 by deletion, substitution or addition of 1 to 10 nucleotides, and being capable of autonomous replication.

[Claim 6]

A replicon-replicating cell, which is prepared by introducing the replicon RNA of any one of claims 1 to 5 into a cell.

[Claim 7]

The replicon-replicating cell of claim 6, wherein the cell is a eukaryotic cell.

[Claim 8]

The replicanting cell of claim 7, wherein the eukaryotic cell is a human liver-derived cell.

[Claim 9]

The replicon-replicating cell of claim 7, wherein the eukaryotic cell is any one cell selected from the group consisting of an Huh7 cell.

[Claim 10]

The replicon RNA of any one of claims 1 to 5, which is for producing or evaluating a therapeutic agent or a diagnostic agent for treatment of hepatitis C virus infection.

[Claim 11]

The replicon-replicating cell of any one of claims 6 to 9, which is for producing or evaluating a therapeutic agent or a diagnostic agent for treatment of hepatitis C virus infection.

[Claim 12]

The replicon RNA of any one of claims 1 to 5, which is for producing a vaccine against hepatitis C virus infection.

[Claim 13]

The replicanting cell of any one of claims 6 to 9, which is for producing a vaccine against hepatitis C virus infection.

#### [Claim 14]

A method of producing a replicon RNA of hepatitis C virus of genotype 2a, comprising extracting the replicon RNA from the replicon-replicating cell of any one of claims 6 to 9.

#### [Claim 15]

A method of producing a viral protein of hepatitis C virus of genotype 2a, comprising culturing the replicon-replicating cell of any one of claims 6 to 9, and obtaining the viral protein from the resulting culture product.

#### [Claim 16]

A method of screening for a substance promoting or suppressing the replication of hepatitis C virus, comprising culturing the replicon-replicating cell of any one of claims 6 to 9 in the presence of a test substance, and detecting the replication of a replicon RNA in the resulting culture product.

#### [Claim 17]

A method of introducing a mutation that increases the replication efficiency to the replicon RNA of hepatitis C virus of genotype 2a, comprising performing once or more the following: obtaining a replicated replicon RNA from the replicon-replicating cell of any one of claims 6 to 9, and introducing the thus obtained replicated replicon RNA into a cell that is different from the replicon-replicating cell so as to prepare a new replicon-replicating cell.

#### [Claim 18]

The method of claim 17, wherein the replication efficiency increases to become at least two times greater than that of the replicon RNA that is introduced at the beginning into the replicon-replicating cell.

#### [Claim 19]

A method of producing a replicon RNA of hepatitis C virus of genotype 2a having increased replication efficiency, comprising performing once or more the following: obtaining a replicated replicon RNA from the replicon-replicating cell of any one of claims 6 to 9, and introducing the thus obtained replicated replicon

RNA into a cell that is different from the replicon-replicating cell so as to prepare a new replicon-replicating cell; and obtaining a replicated replicon RNA from the finally obtained replicon-replicating cell.

#### [Claim 20]

A method of producing a replicon RNA of hepatitis C virus of genotype 2a having increased replication efficiency, comprising detecting a nucleotide mutation or an amino acid mutation between the replicon RNA that is produced so as to have an increased replication efficiency by the method of claim 19 and the replicon RNA that is introduced at the beginning into the replicon-replicating cell; and introducing the thus detected nucleotide mutation or amino acid mutation into a replicon RNA whose replication efficiency is to be increased.

#### [Claim 21]

A replicon RNA, comprising a nucleotide sequence derived from the nucleotide sequence represented by SEQ ID NO: 1 by at least one mutation selected from the group consisting of the following (a) to (h):

- (a) a mutation from A to G at nucleotide site 7157;
- (b) a mutation from C to U at nucleotide site 4955;
- (c) a mutation from A to G at nucleotide site 4936;
- (d) a mutation from A to G at nucleotide site 5000;
- (e) a mutation from A to G at nucleotide site 7288;
- (f) a mutation from G to U at nucleotide site 5901;
- (g) a mutation from A to U at nucleotide site 6113; and
- (h) a mutation from A to G at nucleotide site 2890.

[Detailed Description of Invention]

[0001]

#### [Technical Field of Invention]

The present invention relates to a replicon RNA of the hepatitis C virus of genotype 2a, a replicon-replicating cell wherein the replicon RNA is introduced, and a method of increasing the replication efficiency of the replicon RNA.

#### [0002]

#### [Conventional Art]

The hepatitis C virus (HCV) is a virus belonging to the family Flaviviridae. It has a single-stranded (+) strand sense RNA as its genome and is known to cause hepatitis C. Recent studies have revealed that Hepatitis C virus is classified into a number of types based on genotypes or serotypes. According to the phylogenetic analysis of Simmonds et al., using the nucleotide sequences of the HCV strains, which is currently a mainstream method of classifying HCV genotypes, HCV is classified into 6 genotypes: genotype 1a, genotype 1b, genotype 2a, genotype 2b, genotype 3a and genotype 3b (see Non Patent Literature 1). Each of these types is further classified into several subtypes. The nucleotide sequences of the full-length genomes of a several number of genotypes of HCV have been determined to date (see Patent Literature 1 and Non Patent Literatures 2-5).

[0003]

HCV causes chronic hepatitis by persistent infection. Currently, the main cause of chronic hepatitis observed worldwide is persistent HCV infection. Actually, around 50% of individuals with persistent infection develop chronic hepatitis. Chronic hepatitis in approximately 20% of these patients shifts to liver cirrhosis over the course of 10 to 20 years, and some of these patients further go on to advanced lethal pathological conditions such as hepatic cancer.

Hepatitis C is currently treated mainly by a therapy using interferon- $\alpha$  or interferon- $\beta$ , or a therapy using in combination interferon- $\alpha$  and ribavirin, the purine-nucleoside derivative. However, even when these therapies are performed, the therapeutic effects are observed in only approximately 60% of all the treated patients. When the therapies are ceased after the exertion of the effects, the disease recrudesces in more than half of the patients. The therapeutic effect of interferones is known to relate to HCV genotypes, and is said to be lower

against genotype 1b and higher against genotype 2a (see Non Patent Literature 6).
[0005]

It is an important goal to develop therapeutic agents or prophylactic agents effective against hepatitis C, the incidence rate of which is high in industrial countries, for which currently no causal treatment are present, and which finally bring about serious results. Hence, the development of HCV-specific chemotherapies and vaccine therapies are earnestly desired. A target for the development of an anti-HCV agent may be the suppression of HCV replication or the suppression of infection of cells with HCV.

[0006]

Until recently, propagation of HCV in a cell culture system and infecting cultured cells with HCV have been difficult. Moreover, a chimpanzee has been the only animal that can be infected with HCV and can be used in experiments, so that it has been difficult to carry out studies on the replication mechanism of HCV and the infection mechanism of HCV. However, recently, HCV subgenomic RNA replicons have been prepared as HCV-derived autonomously replicable RNA (see Patent Literature 2 and Non Patent Literatures 7-10), which enables the analysis of the replication mechanism of HCV using cultured cells. These HCV subgenomic RNA replicons are each prepared by substituting structural proteins existing downstream of HCV IRES in the 5' untranslated region of the HCV genomic RNA of genotype 1b with a neomycin resistance gene and EMCV IRES that has been ligated downstream of the resistance gene. It has been demonstrated that this RNA replicon is autonomously replicated in human hepatic cancer cells, Huh7 cells, when introduced into the Huh7 cells followed by culture in the presence of neomycin.

[0007]

However, regarding such intracellular RNA replication systems for HCV, only those using HCV genomic RNA of genotype 1b have been prepared so far. Since there has been a report that different genotypes of HCV differ also in viral

proteins encoded, it may be difficult to sufficiently elucidate the replication mechanism of HCV only by analyzing the subgenomic RNA replicons derived from HCV of genotype 1b. Furthermore, based on the fact that the therapeutic effects of interferons differ depending on the HCV genotypes, it may be particularly difficult to develop an anti-HCV agent having an effect on various types of HCV by the use of only an HCV replication system containing the subgenomic RNA replicon of HCV of genotype 1b.

[8000]

[Patent Literature 1] JP Patent Publication (Kokai) No. 2002-171978 A

[Patent Literature 2] JP Patent Publication (Kokai) No. 2001-17187 A

[Non Patent Literature 1] Simmonds, P. et al, Hepatology, (1994) 10, pp. 1321-1324

[Non Patent Literature 2] Choo et al., Science, (1989) 244, pp. 359-362 [Non Patent Literature 3] Kato et al., J. Med. Virol., (2001) 64(3) pp. 334-339 [Non Patent Literature 4] Okamoto, H et al, J. Gen. Virol., (1992) 73 pp. 673-679 [Non Patent Literature 5] Mori, S. et al, Biochem. Biophis. Res. Commun., (1992) 183, pp. 334-342

[Non Patent Literature 6] Yoshioka et al., Hepatology, (1992) 16(2): pp. 293-299
[Non Patent Literature 7] Lohmann et al., Science, (1999) 285, pp. 110-113
[Non Patent Literature 8] Blight et al., Science, (2000) 290, pp. 1972-1974
[Non Patent Literature 9] Friebe et al., J. Virol., (2001) 75(24): pp. 12047-12057
[Non Patent Literature 10] Ikeda et al., J. Virol., (2002) 76(6): pp. 2997-3006
[0009]

[Problem to be Solved by Invention]

An object of the present invention is to provide an HCV-derived replicon RNA of a HCV genotype for which replicon RNA has not yet been prepared.

[0010]

[Means for Solving the Problem]

As a result of intensive studies to achieve the above object, we have

succeeded in preparing the replicon RNA of HCV genotype 2a. [0011]

That is, the present invention is as follows.

- [1] A replicon RNA, comprising a nucleotide sequence containing at least the 5' untranslated region, the sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein and the 3' untranslated region on the genomic RNA of hepatitis C virus of genotype 2a. Preferably, this replicon RNA further contains at least one selection marker gene or a reporter gene, and at least one IRES sequence.
- [2] A replicon RNA, comprising a nucleotide sequence containing the 5' untranslated region comprising the sequence represented by either SEQ ID NO: 9 or 10; at least one selection marker gene or a reporter gene; an IRES sequence; the sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein on the genomic RNA of hepatitis C virus of genotype 2a; and the 3' untranslated region comprising the nucleotide sequence represented by either SEQ ID NO: 11 or 12.
- [3] The replicon RNA of [1] or [2] above, wherein the genomic RNA of hepatitis C virus of genotype 2a is an RNA comprising the nucleotide sequence represented by SEQ ID NO: 3 or 5.
- [4] A replicon RNA, comprising the following RNA (a) or (b):
- (a) an RNA comprising the nucleotide sequence represented by SEQ ID NO: 1 or 2; and
- (b) an RNA comprising a nucleotide sequence derived from the nucleotide sequence represented by SEQ ID NO: 1 or 2 by deletion, substitution or addition of 1 to 10 nucleotides, and being capable of autonomous replication.
- [5] A replicon-replicating cell, which is prepared by introducing the replicon RNA of any one of [1] to [4] above into a cell. For this replicon-replicating cell, a cell into which the replicon RNA is introduced is preferably a eukaryotic cell, more preferably a human liver-derived cell, and further more preferably an Huh7 cell.

- [6] The replicon RNA of [1] to [4] above, which is for producing or evaluating a therapeutic agent or a diagnostic agent for treatment of hepatitis C virus infection.
- [7] The replican-replicating cell of [5] above, which is for producing or evaluating a therapeutic agent or a diagnostic agent for treatment of hepatitis C virus infection.
- [8] The replicon RNA of [1] to [4] above, which is for producing a vaccine against hepatitis C virus infection.
- [9] The replicon-replicating cell of [5] above, which is for producing a vaccine against hepatitis C virus infection.
- [10] A method of producing a replicon RNA of hepatitis C virus of genotype 2a, comprising extracting the replicon RNA from the replicon-replicating cell of [5] above.
- [11] A method of producing a viral protein of hepatitis C virus of genotype 2a, comprising culturing the replicon-replicating cell of [5] above, and obtaining the viral protein from the resulting culture product.
- [12] A method of screening for a substance promoting or suppressing the replication of hepatitis C virus, comprising culturing the replicon-replicating cell of [5] above in the presence of a test substance, and detecting the replication of a replicon RNA in the resulting culture product.
- [13] A method of introducing a mutation that increases the replication efficiency to the replicon RNA of hepatitis C virus of genotype 2a, comprising performing once or more the following: obtaining a replicated replicon RNA from the replicon-replicating cell of [5] above, and introducing the thus obtained replicated replicon RNA into a cell that is different from the replicon-replicating cell so as to prepare a new replicon-replicating cell. In this method, it is more preferred that the replication efficiency increases to become preferably at least two times greater than that of the replicon RNA that is introduced at the beginning into the replicon-replicating cell.
- [14] A method of producing a replicon RNA of hepatitis C virus of genotype 2a

having increased replication efficiency, comprising performing once or more the following: obtaining a replicated replicon RNA from the replicon-replicating cell of [5] above, and introducing the thus obtained replicated replicon RNA into a cell that is different from the replicon-replicating cell so as to prepare a new replicon-replicating cell; and obtaining a replicated replicon RNA from the finally obtained replicon-replicating cell.

- [15] A method of producing a replicon RNA of hepatitis C virus of genotype 2a having increased replication efficiency, comprising detecting a nucleotide mutation or an amino acid mutation between the replicon RNA that is produced so as to have an increased replication efficiency by the method of [14] above and the replicon RNA that is introduced at the beginning into the replicon-replicating cell; and introducing the thus detected nucleotide mutation or amino acid mutation into a replicon RNA whose replication efficiency is to be increased.
- [16] A replicon RNA, comprising a nucleotide sequence derived from the nucleotide sequence represented by SEQ ID NO: I by at least one mutation selected from the group consisting of the following (a) to (h):
- (a) a mutation from A to G at nucleotide site 7157;
- (b) a mutation from C to U at nucleotide site 4955;
- (c) a mutation from A to G at nucleotide site 4936;
- (d) a mutation from A to G at nucleotide site 5000;
- (e) a mutation from A to G at nucleotide site 7288;
- (f) a mutation from G to U at nucleotide site 5901;
- (g) a mutation from A to U at nucleotide site 6113; and
- (h) a mutation from A to G at nucleotide site 2890.

[0012]

[Mode for Carrying out Invention]

The present invention is explained in detail as follows.

1. HCV-derived replicon RNA according to the present invention

The genome of hepatitis C virus (HCV) is a single-stranded (+) strand

RNA comprising approximately 9600 nucleotides. This genomic RNA comprises the 5' untranslated region (also denoted as 5' NTR or 5' UTR), a translated region composed of a structural region and a non-structural region and the 3' untranslated region (also denoted as 3' NTR or 3' UTR). HCV structural proteins are encoded in the structural region, and a plurality of non-structural proteins are encoded in the non-structural region.

[0013]

Such HCV structural proteins and non-structural proteins are generated through the translation into a continuous form thereof, a polyprotein, from the translated region, restricted degradation of the polyprotein by protease, and then the release of the structural proteins (Core, E1 and E2) and non-structural proteins (NS2, NS3, NS4A, NS4B, NS5A and NS5B), respectively. Among these structural proteins and non-structural proteins, that is, viral proteins of HCV, Core is a core protein, E1 and E2 are envelope proteins, and non-structural proteins (NS2, NS3, NS4A, NS4B, NS5A and NS5B) are proteins involved in virus's own replication. NS2 is known to have metalloprotease activity, and NS3 is known to have serine protease activity (at one-third of the N terminal side) and helicase activity (at two-thirds of the C-terminal side). Furthermore, NS4A is a cofactor for protease activity of NS3, and NS5B has been reported to have RNA-dependent RNA polymerase activity. Furthermore, the genome of HCV of genotype 2a has already been reported to have a similar gene structure (see Patent Literature 1). [0014]

We have constructed RNA capable of autonomous replication using such HCV genome of genotype 2a. Specifically, the HCV-derived replicon RNA of the present invention is an RNA construct, which contains the whole or partial RNA of the HCV genome of genotype 2a and is capable of autonomous replication.

[0015]

In this specification, RNA that is prepared by altering the viral genome of HCV and is capable of autonomous replication is referred to as "replicon RNA" or

"RNA replicon." RNA that is artificially prepared from HCV of genotype 2a and is capable of autonomous replication is referred to as "replicon RNA derived from HCV of genotype 2a." In this specification, the HCV-derived replicon RNA is also referred to as an HCV-RNA replicon.

[0016]

In the present invention, "hepatitis C virus of genotype 2a" or "HCV of genotype 2a" means hepatitis C virus identified as genotype 2a according to the international classification of Simmonds et al. The "hepatitis C virus of genotype 2a" or the "HCV of genotype 2a" of the present invention encompasses not only a virus having naturally occurring HCV genomic RNA, but also a virus having genomic RNA prepared by artificially altering a naturally occurring HCV genomic sequence. Specific examples of HCV of genotype 2a include viruses of JFH-1 strain and the JCH-1 strain (see Patent Literature 1).

Furthermore, "the genomic RNA of hepatitis C virus of genotype 2a" means RNA that comprises the single-stranded (+) strand sense RNA of hepatitis C virus of genotype 2a and has the nucleotide sequence throughout the entire region of its genome. The genomic RNA of hepatitis C virus of genotype 2a is preferably RNA comprising the nucleotide sequence represented by SEQ ID NO: 3 or 5, but is not limited thereto.

[8100]

In the specification of the present application, "5' untranslated region" (5'NTR or 5'UTR), "a sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein," "a sequence encoding Core protein" (Core region or C region), "a sequence encoding E1 protein" (E1 region), "a sequence encoding E2 protein" (E2 region), "a sequence encoding N2 protein" (NS2 region), "a sequence encoding NS3 protein" (NS3 region), "a sequence encoding NS4A protein" (NS4A region), "a sequence encoding NS4B protein" (NS4B region), "a sequence encoding NS5A protein" (NS5A region), "a sequence

encoding NS5B protein" (NS5B region) and "3' untranslated region" (3' NTR or 3' UTR), and other specific regions or sites are determined based on the nucleotide sequence of SEQ ID NO: 3 of the full-length cDNA (JFH-1 clone) encoding the entire region of the genome of the JFH-1 strain, which is HCV of genotype 2a. The nucleotide sequence of SEQ ID NO: 3 can be obtained from the International DNA Data Bank (DDBJ/EMBL/GenBank) by referring to the accession No. AB047639. Specifically, when a particular HCV RNA sequence is aligned with the nucleotide sequence represented by SEQ ID NO: 3, a sequence to be aligned with nucleotides I to 340 on the nucleotide sequence represented by SEQ ID NO: 3 is "5' untranslated region" of the RNA, a sequence to be aligned with the nucleotides 3431 to 9442 on the same are a sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein, a sequence to be aligned with the nucleotides 3431 to 5323 on the same is "a sequence encoding NS3 protein," a sequence to be aligned with the nucleotides 5324 to 5485 on the same is "a sequence encoding NS4A protein," a sequence to be aligned with the nucleotides 5486 to 6268 on the same is a sequence encoding NS4B protein," a sequence to be aligned with the nucleotides 6269 to 7666 on the same is "a sequence encoding NS5A protein," a sequence to be aligned with the nucleotides 7667 to 9442 on the same is "a sequence encoding NS5B protein," and a sequence to be aligned with the nucleotides 9443 to 9678 on the same is "3' untranslated region." Furthermore, in this case, gaps, additions, deletions, substitutions or the like may be present in the "aligned" sequences. Furthermore, the above "particular HCV" is not limited thereto, and includes the JFH-1 strain or JCH-1 strain, or viral strains that are derivatives thereof.

One embodiment of the HCV RNA-replicon according to the present invention is a replicon RNA comprising a nucleotide sequence containing at least the 5' untranslated region, a sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein, and the 3' untranslated region on the

[0019]

genomic RNA of hepatitis C virus of genotype 2a. The replicon RNA may further contain at least one selection marker gene or one reporter gene, and at least one IRES sequence. Furthermore, this replicon RNA may also contain a sequence encoding a viral protein other than NS3, NS4A, NS4B, NS5A and NS5B proteins on the genomic RNA of hepatitis C virus of genotype 2a. [0020]

Another preferred embodiment of HCV RNA-replicon according to the present invention is a replicon RNA comprising a nucleotide sequence containing the 5' untranslated region comprising the nucleotide sequence represented by SEQ ID NO: 9 or 10, at least one selection marker gene or reporter gene, the IRES sequence, a sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein on the genomic RNA of hepatitis C virus of genotype 2a, and the 3' untranslated region comprising the nucleotide sequence represented by SEQ ID NO: 11 or 12. In this case the nucleotide sequences represented by SEQ ID NO: 9 and 10 are sequences of the 5' untranslated regions of rSGREP-JFH1 (SEQ ID NO: 2), respectively, which are replicon RNAs according to the present invention. Furthermore, the nucleotide sequences represented by SEQ ID NO: 11 and 12 are sequences of the 3' untranslated regions of rSGREP-JFH1 (SEQ ID NO: 1) and rSGREP-JCH1 (SEQ ID NO: 2), respectively, which are replicon RNAs according to the present invention.

[0021]

A more preferred embodiment of HCV RNA-replicon according to the present invention is a replicon RNA comprised of an RNA comprising the nucleotide sequence represented by SEQ ID NO: 1 or 2. Furthermore, a replicon RNA comprising a nucleotide sequence derived from the nucleotide sequence represented by SEQ ID NO: 1 or 2 by deletion, substitution or addition of 1 to 50, 1 to 30, 1 to 10, 1 to 6, or 1 to several (2 to 5) nucleotides, and being capable of autonomous replication is also included in the scope of the present invention as a

preferred embodiment. In the present invention, "capable of autonomous replication" means that when replicon RNA is introduced into a cell, the replicon RNA allows its own full-length sequence to be replicated within the cell. For example, this ability of autonomous replication can be confirmed by transfecting replicon RNA into Huh7 cells, culturing the Huh7 cells, extracting RNA from the cells in the thus resulting culture product and conducting Northern blot hybridization for the extracted RNA using a probe that can specifically detect the transfected replicon RNA so as to detect the presence of the replicon RNA. However, examples of such a method are not limited thereto. Specific procedures for confirming the ability of autonomous replication can be conducted according to descriptions given in the Examples of this specification such as those for measuring the ability of colony formation, those for confirming the expression of HCV proteins or those for detecting replicon RNA.

In the present invention, a "selection marker gene" means a gene that can provide a cell with selectivity such that only the cell expressing the gene is selected. A general example of a selection marker gene is an antibiotic resistance gene. In the present invention, preferred examples of a selection marker gene include a neomycin resistance gene, a thymidine kinase gene, a kanamycin resistance gene, a pyrithiamine resistance gene, an adenylyl transferase gene, a Zeocin resistance gene and a puromycin resistance gene. The neomycin resistance gene and the thymidine kinase gene are preferred, and the neomycin resistance gene is more preferred. However, the selection marker gene in the present invention is not limited to these genes.

[0022]

[0023]

Furthermore in the present invention, a "reporter gene" means a marker gene encoding a gene product that is a marker for the expression of the gene. General examples of a reporter gene include structural genes of enzymes that catalyze light emitting reaction or color reaction. Preferred examples of the

reporter gene in the present invention include a transposon Tn9-derived chloramphenical acetyltransferase gene, an Escherichia coli-derived  $\beta$  glucuronidase or  $\beta$  galactosidase gene, a luciferase gene, a green fluorescence protein gene, an acquarin gene from jellyfish, and a secreted placental alkaline phosphatase (SEAP) gene. However, the reporter gene in the present invention is not limited to these genes.

[0024]

Either only one or both of the above selection marker gene and reporter gene may be contained in replicon RNA.

[0025]

In the present invention, "IRES sequence" means an internal ribosome entry site that allows translation to be initiated by binding ribosomes within the inside of RNA. Preferred examples of IRES sequence in the present invention include, but are not limited to, EMCV IRES (the internal ribosome entry site of encephalomyocarditis virus), FMDV IRES and HCV IRES. EMCV IRES and HCV IRES are more preferred, and EMCV IRES is the most preferred sequence.

The replicon RNA according to the present invention may further contain a sequence on the genomic RNA of another HCV strain or HCV of another genotype. For example, the replicon RNA may also contain a fragment of HCV genome of genotype 1b. Examples of another HCV strain include, but are not limited to, HCV-1, HCV-H, HC-J1, HCT-18, H77, DK-7, US11, S14, HCT23, HCV-Th, DR1, DR4, HCT27, S18, SW1, DK9, H90, TD-6E1, S9, HCV-BK, T10, DK1, HC-J4, HCV-J, HK3, HK8, HK5, HCV-G3, IND5, IND8, P10, D1, D3, SW2, T3, S45, SA10, US6, HCV-JK1, HCV-JK4, HCV-JK3, HCV-JK2, HCV-JT, HC-J2, HCV-T, HK4, HC-G9, Z1, Bi, S. I., Cho, J.M., HCV-J6, T4, T9, US10, HC-J5, T2, HC-J7, DK11, SW3, DK8, T8, HC-J8, S83, HK2, HC-J6, HC-J8, BEBE1, HCV-J6, HCV-J8, HD10-2, BR36-9, S52, S54, S2, BR33-1, HK10, DK12, HCV-TR, BA-1, BA-2, DK13, Z1, Z4, Z6, Z7, HK2, SA1, SA4, SA5, SA7, SA13, SA6, NZL1, SA30, EG-

13, HCV-K3a/650, ED43, EUH1480, EUHK2, Th580, VN235, VN405, VN004, JK049, JK046, JFH-1, JCH-1, JCH-2, JCH-3, JCH-4, JCH-5, JCH-6, J6CF and H77.

[0027]

The replicon RNA according to the present invention preferably has the 5' untranslated region on the genomic RNA of HCV of genotype 2a on the 5'-most side, and the 3 'untranslated region on the genomic RNA of HCV of genotype 2a on the 3'-most side. A selection marker gene or a reporter gene may be ligated upstream of the IRES sequence, or upstream or downstream of "the sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein," or inserted in the middle of "the sequence encoding NS3 protein, NS4A protein, NS5B protein."

The replicon RNA according to the present invention more preferably has the 5' untranslated region on the genomic RNA of HCV of genotype 2a on the 5'-most side, and a selection marker gene or a reporter gene, the IRES sequence and "the sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein" downstream of the 5' untranslated region in this order, and the 3' untranslated region on the genomic RNA of HCV of genotype 2a on the 3'-most side.

[0029]

Examples of the replicon RNA according to the present invention may include an RNA containing any foreign gene to be expressed within a cell into which the replicon RNA is introduced, in addition to the sequences as described above. A foreign gene may also be ligated downstream of the 5' untranslated region, or ligated upstream or downstream of a selection marker gene or a reporter gene, or ligated upstream or downstream of "the sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein," or may be inserted in the middle of "the sequence encoding NS3 protein, NS4A protein, NS4B protein,

NS5A protein and NS5B protein." A replicon RNA containing a foreign gene can express a protein encoded by the foreign gene when it is translated within a cell into which the RNA is introduced. Thus, the replicon RNA containing a foreign gene can be appropriately used also for gene therapy or the like, the purpose of which is to generate a particular gene product within a cell.

[0030]

The replicon RNA according to the present invention may further contain a ribozyme. A ribozyme is inserted to ligate a selection marker gene, a reporter gene or a foreign gene on the 5' side in the replicon RNA to those located on the 3' side thereof including the IRES sequence and "the sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein," so that it enables cleavage and separation of the two by the self-cleavage activity of the ribozyme.

In the replicon RNA according to the present invention, the above described selection marker gene, reporter gene, sequences encoding viral proteins on the genomic RNA of hepatitis C virus of genotype 2a, sequences encoding viral proteins of HCV of a genotype other than genotype 2a, a foreign gene or the like are ligated so that they are translated from the replicon RNA in the correct reading frame. Among these sequences, the protein-coding sequences may be ligated to each other via a protease cleavage site and the like, so that after the proteins are expressed as a fusion protein with the polyprotein that is translated from "the sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein" of hepatitis C virus of genotype 2a, the fusion protein is separated by protease into each protein.

# [0032]

### 2. Preparation of replicon RNA according to the present invention

The HCV RNA-replicon according to the present invention can be prepared using any genetic engineering techniques known by persons skilled in the art. The HCV RNA-replicon can be prepared by, for example, the following method,

but the method of preparation is not limited thereto.
[0033]

[0035]

First, DNA corresponding to the entire region of the genomic RNA of hepatitis C virus of genotype 2a is ligated downstream of an RNA promoter according to a standard procedure so as to prepare a DNA clone. As used herein, "DNA corresponding to RNA" means a DNA having a nucleotide sequence derived from the nucleotide sequence of the RNA by substituting U (uracil) with T (thymine). The above RNA promoter is preferably an RNA promoter contained in a plasmid clone. An example of an RNA promoter is not limited, but T7 RNA promoter is particularly preferred.

Next, for the thus prepared DNA clone, for example, the structural region (Core sequence, E1 sequence and E2 sequence) located downstream of the 5' untranslated region and the sequence encoding NS2 protein are substituted with a DNA fragment containing a selection marker gene or a reporter gene and the IRES sequence ligated downstream thereof. In this substitution, portions other than the structural region, such as a fragment on the 3' terminal side of the 5' untranslated region or a part of the sequence encoding NS3 protein may be substituted with a sequence derived from HCV of another genotype.

Subsequently, using the DNA clone after the substitution as a template, RNA is synthesized using RNA polymerase. RNA synthesis can be initiated by a standard procedure from the 5' untranslated region and the IRES sequence. When a template DNA is a plasmid clone, the above DNA region ligated downstream of an RNA promoter is excised by a restriction enzyme from the plasmid clone, and then RNA can be synthesized using the DNA fragment as a template. In addition, preferably the 3' terminus of RNA to be synthesized agrees with the 3' untranslated region of the viral genomic RNA, and no other sequences are added or deleted. The thus synthesized RNA is the replicon RNA according to the present invention.

[0036]

# 3. Preparation of replicon-replicating cells into which replicon RNA from HCV of genotype 2a is introduced

The replicon RNA that is prepared as described above is introduced into cells in which the replicon RNA should be replicated, so that cells wherein the replicon RNA is continuously replicated can be obtained. In this specification, a cell wherein replicon RNA is continuously replicated is referred to as a "replicon-replicating cell."

[0037]

[0038]

As a cell into which replicon RNA is introduced, any cell can be used, as long as it can be subcultured. Such a cell is preferably a eukaryotic cell, more preferably a human liver-derived cell, and further preferably Huh7 cells. As these cells, commercially available cells may be utilized, these cells may be obtained from cell depositories, or cell lines established from any cells (e.g., cancer cells or stem cells) may also be used.

Introduction of replicon RNA into cells can be performed using any technique known by persons skilled in the art. Examples of such an introduction method include electroporation, a particle gun method, a lipofection method, a calcium phosphate method, a microinjection method and a DEAE sepharose method. The method using electroporation is particularly preferred.

[0039]

A replicon RNA of interest may be introduced alone, or may be introduced after it is mixed with other nucleic acids. To vary the quantity of replicon RNA while keeping RNA quantity to be introduced at a certain level, the replicon RNA of interest is mixed with total cellular RNA extracted from cells into which the RNA is introduced, and then the mixture is used for introduction into cells. The quantity of replicon RNA to be used for introduction into cells may be determined depending on the introduction method employed, and is preferably between 1

picogram and 100 micrograms, and more preferably between 10 picograms and 10 micrograms.

[0040]

When replicon RNA containing a selection marker gene or a reporter gene is used for introduction into cells, cells wherein the replicon RNA is introduced and continuously replicated can be selected utilizing the expression of the selection marker gene or the reporter gene. Specifically, for example, such cells into which replicon RNA has been introduced may be cultured in media whereby the cells can be selected by the expression of the selection marker gene or the reporter gene. As an example, when replicon RNA contains a neomycin resistance gene as a selection marker gene, cells into which replicon RNA has been intracellularly introduced are seeded into a culture dish. After 16 to 24 hours of culture, G418 (neomycin) is added to the culture dish at a concentration of 0.05 milligrams/milliliter to 3.0 milligrams/milliliter. The cells are continuously cultured for preferably 10 days to 40 days and more preferably 14 days to 28 days after seeding, while exchanging the culture solution twice a week. Next, surviving cells are stained with crystal violet, so that cells into which the replicon RNA has been introduced and is being continuously replicated can be selected as formed colonies.

[0041]

Cloned cells can be obtained from the formed colonies by cloning surviving cells by a standard procedure, and then continuing the culture of the cells. The thus obtained cell clone wherein the replicon RNA of interest is continuously replicated is referred to as "a replicon-replicating cell clone" in this specification.

[0042]

Regarding the established cell clone, detection of a replicon RNA that has been replicated from the introduced replicon RNA in the cell clone, confirmation of the presence or the absence of the incorporation of a selection marker gene or a reporter gene in the introduced replicon RNA into a host genomic DNA, and confirmation of the expression of an HCV protein are preferably carried out to confirm the fact that a replicon RNA of interest is actually and continuously replicated.

[0043]

A replicon RNA that has been replicated from the introduced replicon RNA in the cell clone (in this specification, hereinafter conveniently referred to as "replicated replicon RNA") may be detected according to any RNA detection method known by persons skilled in the art. For example, detection can be performed by conducting the Northern hybridization method for total RNA extracted from the cell clone using as a probe a DNA fragment specific to the introduced replicon RNA.

[0044]

Furthermore, the presence or the absence of the incorporation of a selection marker gene or a reporter gene in the introduced replicon RNA into a host genomic DNA can be confirmed by, for example, performing PCR for the host genomic DNA extracted from the cell clone to amplify at least a part of the selection marker gene or the reporter gene, and then confirming the presence or the absence of the amplified product. However, examples of relevant methods are not limited thereto. A cell clone for which the amplified product is confirmed is considered to have a selection marker gene or a reporter gene incorporated in the host genome. Thus, regarding the cell clone, the replicon RNA itself may not be continuously replicated within the cell. In this case, whether or not the replicon RNA is continuously replicated can be confirmed by conducting an experiment to confirm the expression of an HCV protein, as described below.

[0045]

The expression of an HCV protein can be confirmed by, for example, causing an antibody against an HCV protein to be expressed from the introduced replicon RNA and to react with a protein extracted from a cell clone. This

method can be conducted by any protein detection method known by persons skilled in the art. Specifically, for example, a protein sample extracted from the cell clone is blotted onto a nitrocellulose membrane, with which an anti-HCV protein antibody (e.g., an anti-NS3-specific antibody or an antiserum collected from a hepatitis C patient) is reacted, and then the anti-HCV protein antibody is detected. If the HCV protein is detected among proteins extracted from the cell clone, it can be concluded that this cell clone continuously replicate HCV-derived replicon RNA to express the HCV protein.

[0046]

As described above, cell clones confirmed to continuously replicate a replicon RNA of interest (replicon-replicating cell clones) can be obtained. present invention, replicon RNA can be obtained by any Furthermore in the method known by persons skilled in the art, for example, by extracting RNA from the replicon-replicating cell, and then separating replicon RNA from the RNA by an electrophoresis method. The present invention also relates to such a method of producing replicon RNA. Moreover, preferably, the replicon-replicating cell according to the present invention can be used for producing HCV proteins. Persons skilled in the art can obtain HCV proteins from the replicon-replicating cells according to any standard method. Specifically, for example, a viral protein of hepatitis C virus of genotype 2a can be produced by culturing repliconreplicating cells, collecting proteins from the resulting culture product (including cultured cells and culture media) by a standard procedure, and then selectively obtaining viral proteins from the proteins by detection or the like using an anti-HCV protein antibody.

[0047]

Moreover, when the replicon-replicating cell according to the present invention continuously replicates replicon RNA containing a foreign gene, a protein encoded by the foreign gene can be obtained by the expression thereof using the replicon-replicating cell. Specifically, for example, the protein encoded

by a foreign gene can be obtained by culturing replicon-replicating cells, collecting proteins from the resulting culture product (including cultured cells and culture media) by a standard procedure, and then selectively obtaining the protein from among the proteins by detection or the like using an antibody against the protein of interest.

[0048]

# 4. Introduction of mutation that increases replication efficiency into replicon RNA from HCV of genotype 2a

Mutation producing enhancement of replication efficiency frequently takes place in the replicon RNA that is replicated or generated in the replicon-replicating cell (replicated replicon RNA) according to the present invention. Such a mutation may be an adaptive mutation.

[0049]

Utilizing this fact, introduction of a mutation enhancing replication efficiency into the replicon RNA according to the present invention can be promoted in the present invention.

[0050]

Specifically, the step comprising obtaining a first replicated replicon RNA by extraction or the like from a first replicon-replicating cell (preferably, a replicon-replicating cell, wherein the replicon RNA according to the present invention has been introduced), and then re-introducing the first replicated replicon RNA into another cell to prepare a second replicon-replicating cell is performed repeatedly once or more, preferably I to 10 times, more preferably I to 5 times, and further preferably I to 2 times, so that the mutation increasing replication efficiency can be introduced at a high frequency into the replicon RNA within the replicon-replicating cells.

[0051]

As a cell into which a replicated replicon RNA is re-introduced, any cell can be used. Such a cell is preferably derived from a biological species that is

the same as that of a cell wherein replicon RNA is introduced at the beginning, more preferably derived from the same tissue derived from the same biological species as that of a cell wherein replicon RNA is introduced at the beginning, and further preferably of a cell line that is the same as that for a cell wherein replicon RNA is introduced at the beginning.

Therefore in the present invention, using the above method, replicon RNA wherein the mutation increasing replication efficiency is introduced can be produced. Specifically, the step comprising obtaining a first replicated replicon RNA by extraction or the like from a first replicon-replicating cell (preferably, a replicon-replicating cell, into which the replicon RNA according to the present invention has been introduced), and then re-introducing the first replicated replicon RNA into another cell so as to prepare a second replicon-replicating cell is performed repeatedly once or more, preferably 1 to 10 times, more preferably 1 to 5 times, and further preferably 1 to 2 times. Subsequently, the replicated replicon RNA is obtained by extraction or the like from the replicon-replicating cell finally obtained at the end of the repeated steps, so that replicon RNA with increased replication efficiency can be produced.

In the present invention, the replication efficiency of a replicon RNA can be increased at least 2 times, preferably 10 to 100 times, and more preferably 100 to 10000 times by the above method.

[0054]

[0053]

[0052]

Regarding the replicon RNA that is produced by such a method so as to have increased replication efficiency, the nucleotide sequence is preferably determined by a known method, for example, by obtaining cDNA by reverse transcription PCR and subjecting such cDNA to sequencing. Furthermore, the thus determined nucleotide sequence or the amino acid sequence encoded by the nucleotide sequence is compared with the nucleotide sequence of replicon RNA

that had been introduced at the beginning into cells, so that adaptive mutation can be identified. As adaptive mutation increasing replication efficiency, in particular, nonsynonymous substitution that mutates an amino acid in a viral protein encoded by replicon RNA is preferred.

[0055]

The present invention also provides a method whereby the replicon RNA of hepatitis C virus of genotype 2a having increased replication efficiency can be produced by introducing the thus identified adaptive mutation into replicon RNA, the replication efficiency of which is to be increased, by a standard procedure. [0056]

The replicon RNA that is produced as described above so as to have increased replication efficiency can be used for producing replicon RNA in large quantity within cells that have been used for the method.

[0057]

The replication efficiency of the replicon RNA according to the present invention can be determined by a method known by persons skilled in the art. For example, it can be determined according to the following method. Replicon RNAs are transfected in quantities of 0.0001, 0.0003, 0.001, 0.003, 0.01, 0.03, 0.1, 0.3 and 1.0 micrograms, respectively, into Huh7 cells, selective culture with G418 is performed for 21 days in a method similar to the above experimental techniques, and then the number of colonies formed (number of colonies) is counted. The quantity of RNA introduced is compared with the number of colonies formed to determine the range of the quantity of the replicon RNA introduced, within which colony formation increases in a quantity-dependent manner. The number of colonies formed within the range is divided by the quantity of RNA introduced, and the resulting value is regarded as the colony forming activity per microgram. This equation is as follows.

Colony forming activity [(Colony Forming Unit, or CFU)/microgram] = Number of colonies formed [colony] / quantity of RNA introduced [microgram]

[0058]

The thus calculated colony forming activity is regarded as a value representing the replication efficiency of replicon RNA introduced. Specifically, the higher the colony forming activity, the higher the replication efficiency of the replicon RNA.

[0059]

In addition, the replication efficiency of replicon RNA can also be shown via a colony-forming ability that is represented by the number of copies of the replicon RNA introduced per formed colony. That is, in this case, the ability can be calculated according to the following equation.

Colony forming ability = number of copies of replicon RNA introduced [copy] / number of formed colonies [colony]
[0060]

# 5. Other embodiments of the present invention

The replicon RNA-replicating cell according to the present invention can also be used as a test system for, for example, screening for a substance that promotes or suppresses the replication of hepatitis C virus. Specifically, for example, replicon replicating cells are cultured in the presence of a test substance, replication of the replicon RNA in the resulting culture product is detected, and then whether or not the test substance promotes or suppresses the replication of the replicon RNA is determined, so that a substance that promotes or suppresses the replication of hepatitis C virus can be screened for. In this case, detection of the replication of the replicon RNA in the resulting culture product may be conducted by detecting the quantity of, or the presence or the absence of, the replicon RNA in the RNAs extracted from the replicon RNA-replicating cell, or by detecting the quantity of, or the presence of, HCV protein contained in the proteins in the culture product or in the replicon RNA-replicating cells contained in the culture product.

[0061]

Such a test cell system using the replicon RNA-replicating cells according to the present invention may be aimed at producing or evaluating a therapeutic agent or a diagnostic agent for treating hepatitis C virus infection. Specific examples of such purposes include the following examples.

- (1) Search for a substance suppressing the proliferation of HCV of genotype 2a.

  Examples of such substance include organic chemicals directly or indirectly affecting the proliferation of HCV of genotype 2a, and antisense oligonucleotides directly or indirectly affecting the proliferation of HCV or the translation of HCV directly or indirectly affecting the proliferation of HCV genome of genotype 2a or proteins by hybridizing to a target sequence in the HCV genome of genotype 2a or a complementary strand thereof.
  - (2) Evaluation of various substances having antiviral action in cell culture.

    Examples of the various substances include substances obtained through rational drug design or high throughput screening (e.g., an isolated and purified enzyme) and the like.
  - and the like.

    (3) Identification of a new target for attack for treating patients infected with HCV of genotype 2a. To identify a host cellular protein that plays an important role in proliferation of HCV virus, for example, the replicon-replicating cell according to the present invention can be used.
    - (4) Evaluation of the ability of HCV virus to acquire resistance against a drug or the like and identification of mutation concerning such resistance.

[0063]

The replicon RNA or replicon RNA-replicating cells according to the present invention may be aimed at the following purposes.

- (5) Production of a viral protein as an antigen that can be used for developing, producing and evaluating a diagnostic agent or a therapeutic agent for hepatitis C virus infection.
- (6) Viral genome replication system for producing HCV virus or virus-like particles that can be used for developing, producing and evaluating a diagnostic

agent or a therapeutic agent for hepatitis C virus infection.

- (7) Production of a vaccine antigen that can be used as a vaccine against HCV of genotype 2a.
- (8) Production of hepatic cell-directed genetic vector that is used after the incorporation of a foreign gene therein for gene therapy.

[0064]

[Examples]

The present invention will be described more specifically based on the following examples and drawings. However, the technical scope of the present invention is not limited by these examples.

[0065]

[Example 1] Preparation of replicon RNA

## (A) Construction of expression vector

DNA corresponding to the entire region of viral genome of hepatitis C virus JFH-1 strain (genotype 2a) that had been separated from patients with fulminant hepatic failure was obtained from a JFH-1 clone containing the fulllength genomic cDNA of the virus strain. The DNA was inserted downstream of T7 RNA promoter sequence that had been inserted in pUC19 plasmid. The thus constructed plasmid DNA is hereinafter referred to as pJFH1. Similarly, DNA corresponding to the entire region of viral genome of hepatitis C virus JCH-1 strain (genotype 2a) that had been separated from patients with chronic hepatitis was obtained from a JCH-1 clone containing the full-length genomic cDNA of the The DNA was inserted downstream of the T7 RNA promoter virus strain. sequence that had been inserted in pUC19 plasmid. The thus constructed plasmid DNA is hereinafter referred to as pJCHI. In addition, the preparation of the above JFH1 clone and JCH-1 clone is described in Patent Literature 1 and Non Patent Literature 3. Moreover, the nucleotide sequence of the full-length cDNA of JFH-1 clone was registered at the International DNA Data Bank (DDBJ/EMBL/GenBank) under accession No. AB047639, and the nucleotide

sequence of the full-length cDNA of the JCH-1 clone under accession No. AB047640.

[0066]

The structures of the thus constructed plasmid DNA pJFH1 and pJCH1 are shown in the upper section of Fig. 1. "T7" represents T7 RNA promoter, and "G" represents dGTP inserted upstream of the 5' end of the inserted JFH-1- or JCH-1-derived DNA and downstream of the 3' end of T7 RNA promoter sequence. A region from "5' NTR" to "3' NTR" is DNA corresponding to the entire genomic region of hepatitis C virus.

[0067]

Next, the structural regions and a part of the non-structural regions of plasmid DNA pJFH1 and pJCH1 were substituted with a neomycin resistance gene (neo; also referred to as a neomycin phosphotransferase gene) and EMCV-IRES (internal ribosome entry site of encephalomyocarditis virus), thereby constructing plasmid DNA pSGREP-JFH1 and pSGREP-JCH1, respectively (lower section of Fig. 1). This construction procedure was conducted according to a previous report (Non Patent Literature 7). Specifically, plasmid pJFH1 and pJCH1 were cleaved with restriction enzymes Age I and Cla I, and between the Age I and Cla I restriction sites, the following fragments were inserted to be ligated; a fragment was prepared by binding of a sequence ranging from 5' NTR to Core region derived from pJFH-1 with the neomycin resistance gene derived from pRSV5NEO by PCR amplification and then cleaving it with restriction enzymes Age I and Pme I, and, a fragment was prepared by binding of sequences ranging from EMCV IRES to NS3 region by PCR amplification and then cleaving it with restriction enzymes Pme I and Cla I.

[0068]

Moreover, a mutation that mutates an amino acid motif GDD to GND, corresponding to the active center of RNA polymerase encoded by the NS5B region, was introduced into the NS5B region in pSGREP-JFH1, thereby preparing

a mutant plasmid clone pSGREP-JFH1/GND.
[0069]

Moreover, a mutation that results in the deletion of a sequence of 10 continuous amino acids containing an amino acid motif GDD corresponding to the active center of RNA polymerase encoded by the NS5B region was introduced into the NS5B region in pSGREP-JFH1, thereby preparing a mutant plasmid clone pSGREP-JFH1/dGDD.

[0070]

The above-prepared mutant clones pSGREP-JFH1/GND and pSGREP-JFH1/dGDD cannot express active NS5B protein, which is required for the replication of replicon RNA, because the amino acid sequence of the active site of NS5B protein encoded by these clones has mutated.

[0071]

### (B) Preparation of replicon RNA

To prepare template DNA for use in synthesis of replicon RNA, the above-constructed expression vectors pSGREP-JFH1, pSGREP-JCH1, pSGREP-JFH1/GND and pSGREP-JFH1/dGDD were each cleaved with a restriction enzyme Xba I.

[0072]

Subsequently, 10 to 20 µg each of these Xba I-cleaved fragments was contained in 50 µl of a reaction solution, and then further treated by 30 minutes of incubation at 30°C with 20 U of Mung Bean Nuclease. Mung Bean Nuclease is an enzyme catalyzing a reaction for selectively degrading a single-stranded portion of double-stranded DNA. Generally, when RNA synthesis is performed using directly the above Xba I-cleaved fragment as a template, a replicon RNA having four nucleotides of CUGA, a part of the recognition sequence of Xba I, excessively added to the 3' terminus would be synthesized. Hence, in this example, Xba I-cleaved fragments were treated with Mung Bean Nuclease, so as to remove the four nucleotides of CUGA from the fragments. The solutions

containing Xba I-cleaved fragments, which had been treated with Mung Bean Nuclease, were treated to remove proteins according to a general method, so that Xba I-cleaved fragments, from which the four nucleotides of CUGA had been removed, were purified and used as template DNAs.

[0073]

Next, from the template DNA, RNA was synthesized in vitro using T7 RNA polymerase. For this RNA synthesis, MEGAscript from Ambion, Inc. was used. Reaction was carried out using 20  $\mu$ l of a reaction solution containing 0.5 to 1.0 micrograms of the template DNA according to the instructions of the manufacturer.

[0074]

After completion of RNA synthesis, DNase (2 U) was added to the reaction solution to conduct reaction at 37°C for 15 minutes. RNA extraction using acidic phenol was further performed to remove the template DNA. RNAs (replicon RNAs) synthesized in this manner from the above template DNAs derived from pSGREP-JFH1, pSGREP-JCH1, pSGREP-JFH1/GND and pSGREP-JFH1/dGDD were respectively named rSGREP-JFH1, rSGREP-JCH1, rSGREP-JFH1/GND and rSGREP-JFH1/dGDD. Regarding the nucleotide sequences of these replicon RNAs, the nucleotide sequence of rSGREP-JFH1 is shown in SEQ ID NO: 1 and Fig. 2, that of rSGREP-JCH1 is shown in SEQ ID NO: 2 and Fig. 3, that of rSGREP-JFH1/GND is shown in SEQ ID NO: 7, and that of rSGREP-JFH1/dGDD is shown in SEQ ID NO: 8.

[0075]

[Example 2] Establishment of replicon-replicating cell clone

(C) Transfection of replicon RNA, determination of colony-forming ability of transfected cells and establishment of cell clones

Each of the above-synthesized replicon RNAs (rSGREP-JFH1, rSGREP-JCH1, rSGREP-JFH1/GND and rSGREP-JFH1/dGDD) was mixed in different quantities with total cellular RNA extracted from Huh7 cells so as to have a total

RNA quantity of 10 µg. Subsequently, the mixed RNA was introduced into Huh7 cells by the electroporation method. The Huh7 cells subjected to the electroporation treatment were seeded into culture dishes, and then cultured for 16 hours to 24 hours. G418 (neomycin) was then added to the culture dishes at different concentrations. Thereafter, culture was continued while exchanging the culture solutions twice a week. After 21 days of culture following seeding, surviving cells were stained with crystal violet. The number of stained colonies was counted, and then the number of colonies obtained per µg of the transfected replicon RNA was calculated.

[0076]

For rSGREP-JFH1 or rSGREP-JCH1-transfected cells, for which colony formation had been observed, colonies of the surviving cells were further cloned from the above culture dishes after 21 days of culture, and were continuously cultured. By such cloning of colonies, several strains of cell clones could be established.

[0077]

For the established cell clones, detection of the replicated replicon RNA, confirmation of the presence or the absence of the incorporation of the neomycin resistance gene into the host genomic DNA, and confirmation of the expression of HCV proteins were performed as described later, in Example 4. Cell clones for which the replication of the replicon had been confirmed in the cells were regarded as replicon-replicating cell clones.

[0078]

#### (D) Colony-forming ability in each transfected cell

As a result of the above transfection, for rSGREP-JFH1-transfected Huh7 cells, the colony-forming ability per μg of the transfected replicon RNA was 94700 CFU (Colony Forming Unit)/μg·RNA when G418 concentration was 1.0 mg/ml (the left column in Fig. 4). In contrast, colony formation was not observed in the Huh7 cells, into which rSGREP-JFH1/dGDD and rSGREP-

JFH1/GND had each been transfected (the central column and the right column in Fig. 4). This suggests that the colony-forming ability confirmed for the Huh7 cells, into which rSGREP-JFH1 replicon RNA had been transfected, depends on the activity of NS5B (RNA polymerase) expressed by rSGREP-JFH1. Specifically, it was considered that in cells that had formed colonies, rSGREP-JFH1 replicon RNA autonomously replicated due to the action of NS5B expressed by rSGREP-JFH1, and the neomycin resistance gene was continuously expressed to maintain G418 resistance, so that cell growth was enabled.

On the other hand, in the Huh7 cells, into which rSGREP-JCH1 had been transfected, no colony formation was observed in the case of 1 to 0.5 mg/ml G418 concentrations (Fig. 5). When G418 concentration was lowered to 0.25 mg/ml, colony formation was observed in the Huh7 cells, into which rSGREP-JCH1 had been transfected as well.

[0800]

[0079]

Furthermore, Xba I-cleaved fragment of the expression vector pSGREP-JFH1 obtained in (B) above was used as a template DNA for RNA synthesis without treating the fragment with Mung Bean Nuclease, so as to synthesize replicon RNA. This replicon RNA was transfected to Huh7 cells in a manner similar to that in (C) above. The replicon RNA that had been prepared without performing Mung Bean Nuclease treatment had the four nucleotides of CUGA excessively added to the 3' terminus.

[0081]

As a result, the colony-forming ability of the Huh7 cells, into which the replicon RNA prepared without treatment with Mung Bean Nuclease had been transfected, decreased to 512 CFU/µg·RNA (the left side in Fig. 6). This result revealed that the sequence on the 3' terminus of the replicon RNA affects the colony-forming ability of the transfected cells.

[0082]

#### [Example 3]

(E) Re-transfection of replicated replicon RNA derived from replicating cells

From the replicon-replicating cell clones that had been established by transfection of rSGREP-JFH1 into Huh7 cells according to descriptions of Example 2, total RNA was extracted by a standard procedure. The number of copies of the replicated replicon RNA contained in the cellular RNA was determined by Northern blot analysis and a quantitative RT-PCR method.

[0083]

Northern blot analysis was performed according to the description in Molecular Cloning, A laboratory Manual, 2<sup>nd</sup> edition, J. Sambrook, E. F. Fritsch, T. Maniatis, Cold Spring Harbor Laboratory Press (1989). Specifically, RNA extracted from the cells was subjected to denaturing agarose electrophoresis. After electrophoresis, the RNA was transferred onto a positively charged nylon membrane. The <sup>32</sup>P-labeled DNA or RNA probe prepared from pSGREP-JFH1 was hybridized to the RNA transferred to the membrane as described above. Next the membrane was washed, and then exposed to a film, so as to detect a replicon-specific RNA band.

[0084]

Detection of the replicon RNA by quantitative RT-PCR was conducted by detecting the 5' untranslated region RNA within HCV RNA according to Takeuchi T, Katsume A, Tanaka T, Abe A, Inoue K, Tsukiyama-Kohara K, Kawaguchi R, Tanaka S and Kohara M., Real-time detection system for quantification of Hepatitis C virus genome, Gastroenterology 116: 636-642 (1999). Specifically, the replicon RNA contained in RNA extracted from the cells was amplified by PCR using synthetic primers: R6-130-S17, 5'-CGGGAGAGCCATAGTGG-3' (SEQ ID NO: 13) and R6-290-R19, 5'-AGTACCACAAGGCCTTTCG-3' (SEQ ID NO: 14); TaqMan Probe; R6-148-S21FT, 5'-CTGCGGAACCGGTGAGTACAC-3' (SEQ ID NO: 15) and an EZ rTth RNA PCR kit, and then detected using an ABI Prism

7700 sequence detector system.

[0085]

Next, aliquots of total cellular RNAs extracted from clone 6 (among the above-mentioned replicon-replicating cell clones) and pool clones (prepared by collecting replicon-replicating cells that had formed colonies from whole one dish and culturing them) were each introduced into another Huh7 cells by retransfection. Total cellular RNA used for the transfection was prepared to contain 1x10<sup>7</sup> copies of replicon RNA based on the number of copies of the above-determined replicon RNA. Transfection was performed as described in (C) above, and then selective culture was performed under G418 concentration conditions of 1 mg/ml. Thus, the colony formation of the replicon-replicating cells was observed (Fig. 7). The colony-forming ability in this case was 1 colony or more per 1x10<sup>6</sup> copies of the replicon RNA used for transfection, when it was calculated from the number of colonies obtained.

[0086]

On the other hand, the number of copies of in vitro synthetic RNA that had been synthesized in vitro using pSGREP-JFH1 as a template and T7 RNA polymerase was approximately  $2x10^{11}$  copies/µg·RNA, when calculated based on the weight and the length of the RNA. The colony-forming ability in the case of using the in vitro synthetic RNA for transfection in a manner similar to the above method was 1 colony per  $5x10^7$  copies. These results revealed that when RNA derived from cells extracted from replicon-replicating cells and in vitro synthetic RNA were each transfected to Huh7 cells as replicon RNA in the same number of copies, the use of the replicon RNA replicated within Huh7 cells resulted in colony-forming ability approximately 50 times higher than that of the in vitro synthetic RNA.

[0087]

[Example 4]

(F) Detection of replicon RNA

According to (E) above, cell clones [clones Nos. 1 to 11] were established by retransfection of total RNA that had been obtained from the replicon-replicating cell clone established by transfection of rSGREP-JFH1 to Huh7 cells to another Huh7 cells. From the established cell clones and pool clones (prepared by collecting cell clones that had formed colonies from whole one dish and then culturing them), respectively, total RNAs were extracted by an acidic phenol extraction method. Subsequently the total RNAs were analyzed by the Northern blot method using a pSGREP-JFH1-specific probe as a probe. As control, total RNA extracted similarly from untransfected Huh7 cells (in Fig. 8, denoted as "Huh7"), a sample prepared by adding 10<sup>7</sup> copies of replicon RNA synthesized in vitro to the total RNA extracted from Huh7 cells (in Fig. 8, denoted as "10<sup>7</sup>"), and a sample (in Fig. 8, denoted as "10<sup>8</sup>") prepared by adding 10<sup>8</sup> copies of replicon RNA synthesized in vitro to the total RNA extracted from Huh7 cells, were used. In Fig. 8, 1 to 11 represent cell clone Numbers.

As a result, RNA of approximately the same size as that of rSGREP-JFH1 was detected using a pSGREP-JFH1-specific probe (Fig. 8). Thus, it was confirmed that the replicon RNA from rSGREP-JFH1 that had been transfected at the beginning replicated and proliferated within the cell clones. In addition, it was shown that the cell clones differed from each other in the quantity of the replicated replicon RNA. In Fig. 8, for example, clones 2, 6, 9 and 10 contained high quantities of the replicated replicon RNA, and clones 4, 8 and 11 contained low quantities of the replicated replicon RNA.

# (G) Confirmation of the presence or the absence of the incorporation of a neomycin resistance gene into genomic DNA

[0089]

For the cell clones that had been obtained by retransfection of replicon RNA as described in Example 3, PCR amplification was performed using neomycin resistance gene-specific primers; sense primer, NEO-S3: 5'-

AACAAGATGGATTGCACGCA-3' (SEQ ID NO: 16) and antisense primer, NEO-R: 5'-CGTCAAGAAGGCGATAGAAG-3' (SEQ ID NO: 17), and the host cellular genomic DNA extracted from each of the cell clones as a template, in order to confirm that the resistance of each of the cell clones against G418 was not due to the incorporation of the neomycin resistance gene into the genome. The cell clones used herein were the cell clones Nos. 1 to 8 obtained by retransfection of rSGREP-JFH1-derived replicated replicon RNA (rSGREP-JFH1-derived cell clones Nos. 1 to 8), and cell clones Nos. 1 to 6 obtained by retransfection of rSGREP-JCH1-derived replicated replicon RNA (rSGREP-JCH1-derived cell clones Nos. 1 to 6). As a result, as shown in Fig. 9, in the eight examined rSGREP-JFH1-derived cell clones, positive clones showing the amplification of the neomycin resistance gene were not observed. For rSGREP-JCH1-derived cell clones, only 1 out of the 6 examined clones was positive (in Fig. 9, lane 3 in the right photograph). It was considered that this positive clone had acquired G418 resistance by the incorporation of the neomycin resistance gene in rSGREP-JCH1derived replicated replicon RNA into the genomic DNA of the host cells. Thus, in the positive clone, unlike other clones, it was thought that the replicon RNA itself did not autonomously replicate within the cells. This was confirmed by the results of the experiment shown in the next (H) that no HCV proteins were detected from the positive clone.

[0090]

#### (H) Detection of HCV protein

Protein was extracted from rSGREP-JFH1- and rSGREP-JCH1-transfected cell clones by a standard procedure, and then analyzed by SDS-PAGE and Western blot method (Fig. 10). The examined cell clones were the same as those used in (G) above: rSGREP-JFH1-derived cell clones Nos. 1 to 8 and rSGREP-JCH1-derived cell clones Nos. 1 to 6. In addition, a cellular extract from the cell obtained by transiently transfecting expression plasmid DNA containing NS3 gene into Huh7 cells was regarded as a positive control (NS3 protein). Furthermore, a

protein extracted from the untransfected Huh7 cells was used as a negative control. A protein sample extracted from each cell clone was blotted onto a PVDF membrane (Immobilon-P, Millipore), and then detection of NS3 protein encoded by replicated replicon RNA was performed using anti-NS3-specific antibody (provided by Dr. Moradpour; Wolk B, et al, J. Virology. 2000, 74: 2293-2304). As shown in Fig. 10, in rSGREP-JFH1-derived cell clones Nos. 1 to 8 and rSGREP-JCH1-derived cell clones Nos. 1, 2 and 4 to 6, proteins of the same size as those of the positive control were detected. In rSGREP-JCH1-derived cell clone No. 3 (the clone detected as a positive clone in (G) above), no expression of NS3 protein was detected. That is, in rSGREP-JCH1-derived cell clone No. 3, no replication of replicon RNA was confirmed. NS3 protein was not detected in the untransfected Huh7 cells, revealing that in cell clones wherein NS3 protein was detected, the transfected replicon RNA autonomously replicated so that NS3 protein was expressed.

[0091]

Moreover, by the use of the serum of a hepatitis C patient as an antibody, the expression of NS5a protein from the replicon RNA was also confirmed in each cell clone for which the expression of NS3 protein had been confirmed as described above.

[0092]

Based on the results of (G) and (H) above, it was confirmed that replicon RNAs were replicated in the cell clones established by transfection of the replicon RNA.

[0093]

[Example 5]

## (I) Analysis of adaptive mutation

According to descriptions of Example 3, total RNA obtained from the replicon-replicating cell clones established through the transfection of rSGREP-JFH1 into Huh7 cells was re-transfected to another Huh7 cells, thereby

establishing 21 cell clones. Total RNA was extracted from each of these cell clones by a standard procedure. cDNA corresponding to the replicon RNA was synthesized using the total RNA as a template, reverse transcriptase Superscript II (Invitrogen) and primer 9641R-IH (5'-GCACTCTCTGCAGTCATGCGGCTCACGGAC-3' (SEQ ID NO: 18)). The composition of a reaction solution for the synthesis of cDNA by reverse transcription reaction is as shown below.

[0094]

Composition of Reaction Solution	Fluid Volume (µl)
5x 1st strand Buffer	4
2 mM dNTP	5
0.1 M DTT	1
9651R-IH primer (100 μM)	1
DW (distilled water)	6.5
Sample RNA (2 mg/mL)	1
RNasin (Promega) (40 U/μL)	0.5
Superscript II RT (Invitrogen)	1
Total	20 μ1
[0095]	

In cDNA synthesis reaction, the above reagents other than RNasin and Superscript II were mixed to prepare a first reaction solution. The solution was heated at 90°C for 3 minutes, and then cooled on ice. Subsequently, RNasin and Superscript II were added to this reaction solution, and then the solution was allowed to react at 42°C for 1 hour, followed by another reaction at 70°C for 15 minutes.

#### [0096]

Furthermore, PCR amplification was performed using the thus obtained cDNA together with five primer sets by the following procedures, so that DNA amplification fragments covering almost all the regions of the replicon RNA were

obtained. The primer sets used and regions amplified by each primer set are shown in Table 1 and Table 2 below.

[0097]

## [Table 1]

Designation of	Primer set		Amplified region
amplified fragment	Primer I	Primer 2	
A/	42S-IH	433R-neo	41 - 470
B/	C/S17ssp	4680R-IH	28 - 3026
C/	4534S-IH	7279R-IH	2880 - 5625
D/	7198S-IH	9367R-IH	5544 - 7713
E/	9247S-NF	9576R-NF	7597 - 7960

# [0098]

In Table 1, an amplified region is represented by nucleotide numbers in rSGREP-JFH1 (SEQ ID NO: 1) that the region corresponds to.

[0099]

## [Table 2]

Primer	Nucleotide sequence (5'→3')	SEQ ID NO:
designation		
42S-IH	CCCCTGTGAGGAACTACTGTCTTCACGC	SEQ ID NO: 19
C/S17ssp	CCGGGAGAGCCATAGTGGTCTGCG	SEQ ID NO: 20
4534S-IH	CCACTCAAAGAAAAAGTGTGACGAGCTCGC	SEQ ID NO: 21
7198S-IH	GGCTTGGGCACGGCCTGA	SEQ ID NO: 22
9247S-NF	GCGGTGAAGACCAAGCTCAAACTCACTCCA	SEQ ID NO: 23
433R-neo	AGAACCTGCGTGCAATCCATC	SEQ ID NO: 24
4680R-IH	CCCGTCATGAGGGCGTCGGTGGC	SEQ ID NO: 25
7279R-IH	ACCAGCAACGGTGGGCGGTTGGTAATC	SEQ ID NO: 26
9367R-RI	GGCACGCGACACGCTGTG	SEQ ID NO: 27
9576R-NF	AGCTAGCCGTGACTAGGGCTAAGATGGAGC	SEQ ID NO: 28

## [0100]

The composition of a reaction solution in this PCR reaction is as follows.
[0101]

Composition of Reaction Solution	Fluid Volume (µl)
Primer 1 (10 μM)	1.0
Primer 2 (10 μM)	1.0
2.5 mM dNTPs	5.0
10x LA Buffer	5.0
MgCl <sub>2</sub> (25 mM)	5.0
LA Taq (TAKARA) (5 U/µl)	0.3
DW (distilled water)	30.7
Template cDNA	2.0
Total	50 μΙ

## [0102]

In addition, PCR reaction conditions are as follows: 95°C for 2 minutes; 35 cycles of 98°C for 10 seconds and then 68°C for 8 minutes; and 72°C for 7 minutes; after which the temperature is kept at 4°C.

#### [0103]

The nucleotide sequence of each PCR product obtained as described above was determined, and then the RNA sequence corresponding to the DNA sequence was compared with the sequence of rSGREP-JFH1. The results are shown in Table 3.

## [0104]

[Table 3]

Region	Synonymous substitution	Nonsynonymous substitution	Total number of mutations
NS3	0	5	5
NS4A	0	2	2
NS4B	0	3	3
NS5A	0	7	7
NS5B	3	5	8
Total	3	22	25

## [0105]

As shown in Table 3, total number of nucleotide mutations observed in 21 cell clones was 25. 22 of these mutations were nonsynonymous substitutions inducing amino acid mutation. Types of these mutations are as shown in Table 4. In addition, the positions of these mutations in the non-structural region are shown in Fig. 11.

[0106] [Table 4]

Clone Mutation site designation Nucleotide No. Nucleotide mutation Amino acid mutation Amino acid No. C1 7098  $A \Rightarrow G$ None  $Y \rightarrow C$ 7157  $A \Rightarrow G$ 2824 C2  $C \Rightarrow U$ 4955  $A \Rightarrow V$ 2090 C34936  $A \Rightarrow G$  $T \Rightarrow A$ 2084  $Y \Rightarrow C$ 5000  $A \Rightarrow G$ 2105  $A \Rightarrow G$ 7287 None 7288  $A \Rightarrow G$  $M \Rightarrow V$ 2868 C4 5901  $G \Rightarrow U$ 2405  $E \Rightarrow D$ 6113  $H \Rightarrow L$  $A \Rightarrow U$ 2476

 $A \Rightarrow G$ 

 $A \Rightarrow G$ 

 $K \Rightarrow E$ 

None

1402

#### [0107]

C5

<u>C6</u>

2890

7209

In Table 4 and Fig. 11, "C1 to C6" represent replicon-replicating cell clones C1 to C6 having replicon RNA found to have mutations. "Nucleotide No." shows the corresponding nucleotide numbers within the nucleotide sequence of replicon RNA rSGREP-JFH1 (SEQ ID NO: 1). "Amino acid No." shows the corresponding amino acid numbers within the amino acid sequence encoded by the JFH-1 clone (SEQ ID NO: 4). The types of nucleotides and amino acids at

mutation sites are described according to their general notations. As shown in Table 4, in clone C2, a nucleotide corresponding to nucleotide No. 4955 of SEQ ID NO: 1 on the replicon RNA mutated from C (cytosine) to U (uracil), which results in a mutation of an amino acid corresponding to amino acid No. 2090 of SEQ ID NO: 4 from A (alanine) to V (valine).

[0108]

Furthermore, mutation positions shown in Fig. 11 are shown with bar lines with the nucleotide numbers shown in Table 4. A thick bar line represents nonsynonymous substitution, and a thin bar line represents synonymous substitution.

[0109]

There were 2 clones having no nucleotide mutations at all that cause amino acid mutations. When Northern blot analysis was conducted for the 2 clones, it was shown that in these 2 clones, the quantity of replicon RNAs replicated was lower than those in the cell clones that had replicated replicon RNAs having a nucleotide mutation that causes an amino acid mutation. Hence, it was considered that the nucleotide mutation causing an amino acid mutation within the replicon RNA was an adaptive mutation for increasing the replication efficiency of the replicon RNA in Huh7 cells.

[0110]

[Effects of Invention]

According to the present invention, an HCV-RNA replicon derived from the genotype 2a strain of HCV was obtained for the first time. The replicon-replicating cell according to the present invention can be used as a culture system for the continuous production of RNA and HCV proteins derived from HCV of genotype 2a. Furthermore, the replicon-replicating cell according to the present invention is useful as a test system for screening for various substances that affect HCV replication and/or the translation of HCV proteins.

[0111]

## [Sequence Listing]

<400> 1

#### SEQUENCE LISTING

```
<110> Toray Industries Inc.
      Tokyo Metropolitan Organization for Medical Research
      Johannes Gutenberg-Universitaet Mainz
<120> Establishment of the genotype 2a Hepatitis C virus subgenomic
replicon
<130> P03-0429
<140>
<141>
<160> 28
<170> Patentîn Ver. 2.1
<210> 1
<211> 8024
<212> RNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: replicon
<220>
<223> Inventor: Wakita, Takaji
     Inventor: Kato, Takanobu
     Inventor: Date, Tomoko
```

45

accugecceu	aauaggggcg	acacucegee	augaaucacu	ccccugugag	gaacuacugu	60
cuucacgcag	aaagegeeua	gecauggegu	uaguaugagu	gucguacagc	cuccaggccc	120
ccccucccg	ggagagccau	aguggucugc	ggaaccggug	aguacaccgg	aauugccggg	180
aagacugggu	ccuuucuugg	auaaacccac	ucuaugeceg	gccauuuggg	egugeeeeeg	240
caagacugcu	agccgaguag	cguuggguug	cgaaaggccu	ugugguacug	ccugauaggg	300
cgcuugcgag	ugccccggga	ggucucguag	accgugcacc	augagcacaa	auccuaaacc	360
ucaaagaaaa	accaaaagaa	acaccaaccg	ucgcccaaug	auugaacaag	auggauugca	420
cgcagguucu	ccggccgcuu	ggguggagag	gcuauucggc	uaugacuggg	cacaacagac	480
aaucggcugc	ucugaugccg	ccguguuccg	gcugucagcg	caggggcgcc	cgguucuuuu	540
ugucaagacc	gaccuguccg	gugeceugaa	ugaacugcag	gacgaggcag	cgcggcuauc	600
guggeuggee	acgacgggcg	uuccuugcgc	agcugugcuc	gaeguuguca	cugaagcggg	660
aagggacugg	cugcuauugg	gcgaagugcc	ggggcaggau	cuccugucau	cucaccuuge	720
uccugcegag	aaaguaucca	ucauggcuga	ugcaaugcgg	cggcugcaua	cgcuugaucc	780
ggcuaccugc	ccauucgacc	accaagcgaa	acaucgcauc	gagcgagcac	guacucggau	840
ggaagceggu	cuugucgauc	aggaugaucu	ggacgaagag	caucaggggc	uegegeeage	900
cgaacuguuc	gccaggcuca	aggcgcgcau	gcccgacggc	gaggaucucg	ucgugaccca	960
uggcgaugcc	ugcuugeega	auaucauggu	ggaaaauggc	cgcuuuucug	gauucaucga	1020
cuguggeegg	cugggugugg	cggaccgcua	ucaggacaua	gcguuggcua	cccgugauau	1080
ugcugaagag	cuuggeggeg	aaugggcuga	ccgcuuccuc	gugcuuuacg	guaucgccgc	1140
ucccgauucg	cagcgcaucg	ccuucuaucg	ccuucuugac	gaguucuucu	gaguuuaaac	1200
ccucucccuc	ccccccccu	aacguuacug	geegaageeg	cuuggaauaa	ggccggugug	1260
eguuugucua	uauguuauuu	uccaccauau	ugccgucuuu	uggcaaugug	agggcccgga	1320
aaccuggccc	ugucuucuug	acgagcauuc	cuaggggucu	uuccccucuc	gccaaaggaa	1380
ugcaaggucu	guugaauguc	gugaaggaag	caguuccucu	ggaagcuucu	ugaagacaaa	1440
caacgucugu	agegacecuu	ugcaggcagc	ggaacccccc	accuggcgac	aggugccucu	1500
gcggccaaaa	gccacgugua	uaagauacac	cugcaaaggc	ggcacaaccc	cagugccacg	1560
uugugaguug	gauaguugug	gaaagaguca	aauggcucuc	cucaagcgua	uucaacaagg	1620
ggcugaagga	ugcccagaag	guaccccauu	guaugggauc	ugaucugggg	ccucggugca	1680
caugcuuuac	auguguuuag	ucgagguuaa	aaaaacgucu	aggeeeceeg	aaccacgggg	1740
acgugguuuu	ccuuugaaaa	acacgaugau	accauggcuc	ccaucacugc	uuaugcccag	1800
caaacacgag	geeueeuggg	cgccauagug	gugaguauga	eggggeguga	caggacagaa	1860
caggccgggg	aaguccaaau	ccuguccaca	gucucucagu	ccuuccucgg	aacaaccauc	1920

ucggggguuu	uguggacugu	uuaccacgga	gcuggcaaca	agacucuage	eggeuuaegg	1980
gguccgguca	cgcagaugua	cucgagugcu	gagggggacu	ugguaggcug	gcccagcccc	2040
ccugggacca	agucuuugga	gccgugcaag	uguggageeg	ucgaccuaua	ucuggucacg	2100
cggaacgcug	augucauccc	ggcucggaga	cgcggggaca	agcggggagc	auugcucucc	2160
ccgagaccca	uuucgaccuu	gaaggggucc	ucgggggggc	cggugcucug	cccuaggggc	2220
cacgucguug	ggcucuuccg	agcagcugug	ugcucucggg	gcguggccaa	auccaucgau	2280
uucauccccg	uugagacacu	cgacguuguu	acaaggucuc	ccacuuucag	ugacaacagc	2340
acgccaccgg	cugugcccca	gaccuaucag	gucggguacu	ugcaugcucc	aacuggcagu	2400
ggaaagagca	ccaagguccc	ugucgcguau	gccgcccagg	gguacaaagu	acuagugcuu	2460
aaccccucgg	uagcugccac	ccugggguuu	ggggcguacc	uauccaaggc	acauggcauc	2520
aaucccaaca	uuaggacugg	agucaggacc	gugaugaccg	gggaggccau	cacguacucc	2580
acauauggca	aauuucucgc	cgaugggggc	ugcgcuagcg	gcgccuauga	caucaucaua	2640
ugcgaugaau	gccacgcugu	ggaugeuace	uccauucucg	gcaucggaac	gguccuugau	2700
caagcagaga	cagccggggu	cagacuaacu	gugcuggcua	eggecacace	ccccggguca	2760
gugacaaccc	cccaucccga	uauagaagag	guaggeeueg	ggcgggaggg	ugagaucece	2820
uucuauggga	gggcgauucc	ccuauccugc	aucaagggag	ggagacaccu	gauuuucugc	2880
cacucaaaga	aaaaguguga	cgagcucgcg	geggeeeuue	ggggcauggg	cuugaaugcc	2940
guggcauacu	auagaggguu	ggacgucucc	auaauaccag	cucagggaga	uguggugguc	3000
guegeeaeeg	acgeceucau	gacgggguac	acuggagacu	uugacuccgu	gaucgacugc	3060
aauguagcgg	ucacccaagc	ugucgacuuc	agccuggacc	ccaccuucac	uauaaccaca	3120
cagacugucc	cacaagacgc	ugucucacge	agucagegee	gcgggcgcac	agguagagga	3180
agacagggca	cuuauaggua	uguuuccacu	ggugaacgag	ccucaggaau	guuugacagu	3240
guagugcuuu	gugagugcua	cgacgcaggg	gcugcguggu	acgaucucac	accageggag	3300
accaccguca	ggcuuagagc	guauuucaac	acgcccggcc	uacccgugug	ucaagaccau	3360
cuugaauuuu	gggaggcagu	uuucacegge	cucacacaca	uagacgccca	cuuccucucc	3420
caaacaaagc	aageggggga	gaacuucgcg	uaccuaguag	ccuaccaagc	uacggugugc	3480
gccagagcca	aggeceeuce	cccguccugg	gacgccaugu	ggaagugccu	ggcccgacuc	3540
aagccuacgc	uugegggeee	cacaccucuc	cuguaccguu	ugggcccuau	uaccaaugag	3600
gucacccuca	cacacccugg	gacgaaguac	aucgccacau	gcaugcaagc	ugaccuugag	3660
gucaugacca	gcacgugggu	ccuagcugga	ggaguccugg	cageeguege	cgcauauugc	3720
cuggegacug	gaugeguuue	caucaucggc	cgcuugcacg	ucaaccagcg	agucgucguu	3780
gegeeggaua	aggagguccu	guaugaggcu	uuugaugaga	uggaggaaug	cgccucuagg	3840

geggeueuca	ucgaagaggg	gcagcggaua	gccgagaugu	ugaaguccaa	gauccaaggc	3900
uugcugcagc	aggccucuaa	gcaggcccag	gacauacaac	ccgcuaugca	ggcuucaugg	3960
cccaaagugg	aacaauuuug	ggccagacac	auguggaacu	ucauuagcgg	cauccaauac	4020
cucgcaggau	ugucaacacu	gccagggaac	cccgcggugg	cuuccaugau	ggcauucagu	4080
gccgcccuca	ccaguccguu	gucgaccagu	accaccaucc	uucucaacau	caugggaggc	4140
ugguuagcgu	cccagaucgc	accacccgcg	ggggccaccg	gcuuugucgu	caguggccug	4200
gugggggcug	ccgugggcag	cauaggccug	gguaaggugc	ugguggacau	ccuggcagga	4260
uauggugcgg	gcauuucggg	ggcccucguc	gcauucaaga	ucaugucugg	cgagaagccc	4320
ucuauggaag	augucaucaa	ucuacugccu	gggauccugu	cuccgggagc	ccugguggug	4380
ggggucaucu	gegeggeeau	ucugegeege	cacgugggac	cgggggaggg	cgcgguccaa	4440
uggaugaaca	ggcuuauugc	cuuugcuucc	agaggaaacc	acguegeece	uacucacuac	4500
gugacggagu	cggaugcguc	gcagcgugug	acccaacuac	uuggcucucu	uacuauaacc	4560
agccuacuca	gaagacucca	caauuggaua	acugaggacu	gececaucee	augcuccgga	4620
uccuggcucc	gcgacgugug	ggacuggguu	ugcaccaucu	ugacagacuu	caaaaauugg	4680
cugaccucua	aauuguuccc	caagcugccc	ggccuccccu	ucaucucuug	ucaaaagggg	4740
uacaagggug	ugugggccgg	cacuggeauc	augaccacgc	gcugcccuug	cggcgccaac	4800
aucucuggca	augucegeeu	gggcucuaug	aggaucacag	ggccuaaaac	cugcaugaac	4860
accuggcagg	ggaccuuucc	uaucaauuge	uacacggagg	gecagugege	gccgaaaccc	4920
cccacgaacu	acaagaccgc	caucuggagg	guggcggccu	cggaguacgc	ggaggugacg	4980
cagcaugggu	cguacuccua	uguaacagga	cugaccacug	acaaucugaa	aauuccuugc	5040
caacuaccuu	cuccagaguu	uuucuccugg	guggacggug	ugcagaucca	uagguuugca	5100
cccacaccaa	agccguuuuu	ccgggaugag	gucucguucu	gcguugggcu	uaauuccuau	5160
gcugucgggu	eccageuucc	cugugaaccu	gagecegaeg	cagacguauu	gagguccaug	5220
cuaacagauc	cgccccacau	cacggcggag	acugeggege	ggcgcuuggc	acggggauca	5280
ccuccaucug	aggegageue	cucagugage	cagcuaucag	caccgucgcu	gcgggccacc	5340
ugcaccaccc	acagcaacac	cuaugacgug	gacauggucg	augecaaceu	gcucauggag	5400
ggeggugugg	cucagacaga	gccugagucc	agggugcccg	uucuggacuu	ucucgagcca	5460
auggccgagg	aagagagcga	ccuugagccc	ucaauaccau	cggagugcau	gcuccccagg	5520
ageggguuue	cacgggccuu	accggcuugg	gcacggccug	acuacaaccc	gccgcucgug	5580
gaaucgugga	ggaggccaga	uuaccaaccg	cccaccguug	cugguugugc	ucucccccc	5640
cccaagaagg	ccccgacgcc	ucccccaagg	agacgccgga	cagugggucu	gagcgagagc	5700
accauaucag	aagcccucca	gcaacuggcc	aucaagaccu	uuggccagcc	ccccucgagc	5760

ggugaugcag	gcucguccac	gggggcgggc	gccgccgaau	ccggcggucc	gacguccccu	5820
ggugagcegg	ccccucaga	gacagguucc	gccuccucua	ugeceeeccu	cgagggggag	5880
ccuggagauc	cggaccugga	gucugaucag	guagagcuuc	aaccuccccc	ccaggggggg	5940
gggguagcuc	ccgguucggg	cucggggucu	uggucuacuu	gcuccgagga	ggacgauacc	6000
accgugugcu	gcuccaugue	auacuccugg	accggggcuc	uaauaacucc	cuguagecee	6060
gaagaggaaa	aguugccaau	caacccuuug	aguaacucgc	uguugcgaua	ccauaacaag	6120
guguacugua	caacaucaaa	gagegeeuca	cagagggcua	aaaagguaac	uuuugacagg	6180
acgcaagugc	ucgacgccca	uuaugacuca	gucuuaaagg	acaucaagcu	ageggeuuee	6240
aaggucagcg	caaggcuccu	caccuuggag	gaggcgugcc	aguugacucc	accccauucu	6300
gcaagaucca	aguauggauu	cggggccaag	gagguccgca	gcuuguccgg	gagggccguu	6360
aaccacauca	aguccgugug	gaaggaccuc	cuggaagacc	cacaaacacc	aauucccaca	6420
accaucaugg	ccaaaaauga	gguguucugc	guggaccccg	ccaagggggg	uaagaaacca	6480
gcucgccuca	ucguuuaccc	ugaccucggc	guccgggucu	gcgagaaaau	ggcccucuau	6540
gacauuacac	aaaagcuucc	ucaggcggua	augggagcuu	ccuauggcuu	ccaguacucc	6600
ccugcccaac	ggguggagua	ucucuugaaa	gcaugggcgg	aaaagaagga	ccccaugggu	6660
uuuucguaug	auacccgaug	cuucgacuca	accgucacug	agagagacau	caggaccgag	6720
gaguccauau	accaggccug	cucccugece	gaggaggccc	gcacugccau	acacucgeug	6780
acugagagac	uuuacguagg	agggcccaug	uucaacagca	agggucaaac	cugcgguuac	6840
agacguugcc	gegeeagegg	ggugcuaacc	acuagcaugg	guaacaccau	cacaugcuau	6900
gugaaagccc	uageggeeug	caaggcugcg	gggauaguug	cgcccacaau	gcugguaugc	6960
ggcgaugacc	uaguagucau	cucagaaagc	caggggacug	aggaggacga	gcggaaccug	7020
agagecuuca	cggaggccau	gaccagguac	ucugececue	cuggugaucc	ccccagaccg	7080
gaauaugacc	uggagcuaau	aacauccugu	uccucaaaug	ugucuguggc	guugggcccg	7140
cggggccgcc	gcagauacua	ccugaccaga	gacccaacca	cuccacuege	ccgggcugcc	7200
ugggaaacag	uuagacacuc	cccuaucaau	ucauggcugg	gaaacaucau	ccaguaugcu	7260
ccaaccauau	ggguucgcau	gguccuaaug	acacacuucu	ucuccauucu	caugguccaa	7320
gacacccugg	accagaaccu	caacuuugag	auguauggau	caguauacuc	cgugaauccu	7380
uuggaccuuc	cagecauaau	ugagagguua	cacgggcuug	acgccuuuuc	uaugcacaca	7440
uacucucacc	acgaacugac	geggguggeu	ucageceuca	gaaaacuugg	ggcgccaccc	7500
cucagggugu	ggaagagucg	ggcucgcgca	gucagggcgu	cccucaucuc	ccguggaggg	7560
aaageggeeg	uuugeggeeg	auaucucuuc	aauugggcgg	ugaagaccaa	gcucaaacuc	7620
acuccauugc	cggaggcgcg	ccuacuggac	uuauccaguu	gguucaccgu	eggegeegge	7680

<210> 2

<211> 8024

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: replicon

<400> 2

 acccgccccu
 aauaggggcg
 acacuccgcc
 augaaucacu
 cccugugag
 gaacuacugu
 120

 cuucacgcag
 aaagcgucua
 gccauggcgu
 uaguaugagu
 gucguacagc
 cuccaggccc
 120

 cccccucccg
 ggaagaccau
 agugucugc
 ggaaccggug
 aguacaccgg
 aauugccgg
 180

 aagacuggu
 ccuuucuugg
 auaaacccac
 ucuaugccg
 gccauuuggg
 cuguacacgg
 240

 caagacugcu
 agccgaguag
 cguuggguug
 cgaaaggccu
 ugugguacug
 ccugauaggg
 300

 ugcuugcgag
 ugcccggga
 ggucucguag
 accgugcac
 augagaacaa
 aucccaaacg
 360

 ucaaagaaa
 accaaaagaa
 acacuaaccg
 ucgcccaaug
 auugaacaag
 auggauugca
 420

 cgcaagguucu
 ccggccgcuu
 gguugaagag
 gcuauucggc
 aaggggcgc
 cgguucuuu
 540

 aaucggcugc
 ucugaugccg
 gugucuaggc
 gacgaggcg
 cgguucuuu
 540

 ggacaggacugg
 uuccuuggg
 agcugugcug
 gacguugucu
 cugaagggg
 660

 aagggacugg

cgaacuguuc	gccaggcuca	aggegegeau	gcccgacggc	gaggaucucg	ucgugaccca	960
uggcgaugcc	ugcuugccga	auaucauggu	ggaaaauggc	cgcuuuucug	gauucaucga	1020
cuguggccgg	cugggugugg	cggaccgcua	ucaggacaua	gcguuggcua	cccgugauau	1080
ugcugaagag	cuuggcggcg	aaugggcuga	ccgcuuccuc	gugcuuuacg	guaucgccgc	1140
ucccgauucg	cagegeaueg	ccuucuaucg	ccuucuugac	gaguucuucu	gaguuuaaac	1200
ccucucccuc	ecccccccu	aacguuacug	gccgaagccg	cuuggaauaa	ggccggugug	1260
cguuugucua	uauguuauuu	uccaccauau	ugccgucuuu	uggcaaugug	agggcccgga	1320
aaccuggccc	ugucuucuug	acgagcauuc	cuaggggucu	uuccccucuc	gccaaaggaa	1380
ugcaaggucu	guugaauguc	gugaaggaag	caguuccucu	ggaagcuucu	ugaagacaaa	1440
caacgucugu	agegaeceuu	ugcaggcagc	ggaacccccc	accuggcgac	aggugeeucu	1500
geggeeaaaa	gccacgugua	uaagauacac	cugcaaaggc	ggcacaaccc	cagugecacg	1560
uugugaguug	gauaguugug	gaaagaguca	aauggcucuc	cucaagcgua	uucaacaagg	1620
ggcugaagga	ugcccagaag	guaccccauu	guaugggauc	ugaucugggg	ccucggugca	1680
caugcuuuac	auguguuuag	ucgagguuaa	aaaaacgucu	aggececeeg	aaccacgggg	1740
acgugguuuu	ccuuugaaaa	acacgauaau	accauggece	ccaucaccgc	uuacgcccag	1800
cagacacgag	gucucuuggg	cucuauagug	gugagcauga	cggggcguga	caagacagaa	1860
caggccgggg	agguccaagu	ccuguccaca	gucacucagu	ccuuccucgg	aacauccauu	1920
ucgggggucu	uauggacugu	uuaccaegga	gcuggcaaca	agacacuage	eggeuegegg	1980
ggcccgguca	cgcagaugua	cucgagegee	gagggggacu	uggucgggug	geccageceu	2040
ccugggacca	aaucuuugga	gccguguacg	uguggagcgg	ucgaccugua	uuuggucacg	2100
cggaacgcug	augucaucce	ggcucgaaga	cgcggggaca	agcggggagc	geugeueuee	2160
ccgagaecce	uuucgaccuu	gaaggggucc	ucggggggac	cugugcuuug	cccuaggggc	2220
cacgcugucg	gaaucuuccg	ggcagcugug	ugcucucggg	guguggcuaa	guccauagau	2280
uucaucceeg	uugagacgcu	cgacaucgúc	acgcggucuc	ccaccuuuag	ugacaacagc	2340
acaccaccag	cugugeeeca	gaccuaucag	gugggguacu	ugcacgcccc	cacuggcagu	2400
ggaaaaagca	ccaagguccc	cgucgcguac	gccgcccagg	gguauaaagu	geuggugeue	2460
aaucccucgg	uggcugccac	ccugggauuu	ggggcguacu	uguccaaggc	acauggcauc	2520
aaccccaaca	uuaggacugg	agucagaacu	gugacgaccg	gggagcccau	uacauacucc	2580
acguauggua	aauuccucgc	cgaugggggc	ugcgcaggcg	gegecuauga	caucaucaua	2640
ugcgaugaau	gccacucugu	ggaugcuacc	acuauucucg	gcaucgggac	aguecuugae	2700
caagcagaga	cagccggggu	caggcuaacu	guacuggcca	cggccacgcc	ccccgggucg	2760
gugacaaccc	cccaucccaa	uauagaggag	guageceueg	gacaggaggg	ugagaucccc	2820

uucuauggga	gggcguuucc	ccugucuuac	aucaagggag	ggaggcacuu	gauuuucugc	2880
cacucaaaga	aaaaguguga	cgagcucgca	acggcccuuc	ggggcauggg	cuugaacgcu	2940
guggcauauu	acagaggguu	ggacgucucc	auaauaccaa	cucaaggaga	uguggugguc	3,000
guugecaceg	acgeceucau	gacgggguau	acuggagacu	uugacuccgu	gaucgacugc	3060
aacguagcgg	ucacccagge	cguagacuuc	agccuggacc	ccaccuucac	uauaaccaca	3120
cagacugucc	cgcaagacgc	ugucucacgu	agucagogoo	gagggcgcac	ggguagagga	3180
agacugggca	uuuauaggua	uguuuccacu	ggugagcgag	ccucaggaau	guuugacagu	3240
guaguacucu	gugagugcua	cgacgcagga	gcugcuuggu	augagcucuc	accaguggag	3300
acgaccguca	ggcucagggc	guauuucaac	acgccuggcu	ugccugugug	ccaggaccac	3360
cuugaguuuu	gggaggcagu	uuucaccggc	cucacacaca	uagacgcuca	uuuccuuucc	3420
cagacaaagc	agucggggga	aaauuucgca	uacuuaguag	ccuaucaggc	cacagugugc	3480
gccagggcca	aagegeeeee	cccguccugg	gacgucaugu	ggaagugcuu	gacucgacuc	3540
aagcccacgc	uugugggccc	uacaccucuc	cuguaccguu	ugggcucugu	uaccaacgag	3600
gucacccuua	cacaccccgu	gacaaaauac	aucgccacau	gcaugcaagc	ugaccucgag	3660
gucaugacca	gcacgugggu	ccuggcuggg	ggagucuuag	cageeguege	cgcguauugc	3720
uuagcgaccg	gguguguuuc	caucauuggc	cguuuacaca	ucaaccagcg	agcugucguc	3780
gcuccggaca	aggagguccu	cuaugaggcu	uuugaugaga	uggaggaaug	ugccuccaga	3840
geggeueuce	uugaagaggg	gcagcggaua	gccgagaugc	ugaaguccaa	gauccaaggc	3900
uuauugcagc	aagccucuaa	acaggcccag	gacauacaac	ccgcugugca	agcuucgugg	3960
cccaagaugg	agcaauucug	ggccaaacau	auguggaacu	ucauaagcgg	cauucaguac	4020
cucgcaggac	ugucaacacu	gccagggaac	ccugcugugg	cuuccaugau	ggcauucagc	4080
gccgcccuca	ccaguccguu	gucaacuagc	accaccaucc	uucuuaacau	ucuggggggc	4140
uggcuggcgu	cccaaauugc	gccacccgcg	ggggccacug	gcuuuguugu	caguggccug	4200
gugggagcug	cuguuggcag	cauaggcuug	gguaaagugc	ugguggacau	ccuggcaggg	4260
uauggugcgg	gcauuucggg	ggcccucguc	gcguuuaaga	ucaugucugg	egagaageee	4320
uccauggagg	augucaucaa	cuugcugccu	gggauucugu	cuccaggugc	ucugguggug	4380
ggagucaucu	gcgcggccau	ucugegeege	caugugggac	cgggggaagg	cgcgguccaa	4440
uggaugaaca	ggcuuaucgc	cuucgcuucc	agaggaaacc	acguegeece	uacucacuac	4500
gugacggagu	cggaugcguc	gcagcguguc	acccaacuge	uuggcucucu	cacuauaacu	4560
agucuacuca	ggagacuuca	caacuggauc	acugaggauu	gccccauccc	augegeegge	4620
ucguggcucc	gcgaugugug	ggacuggguc	uguaccaucc	uaacagacuu	uaagaacugg	4680
cugaccucca	agcuguuccc	aaagaugccu	ggccuccccu	uuaucucuug	ccaaaagggg	4740

uacaagggcg	ugugggccgg	cacuggcauc	augaccacac	gaugeeccug	cggcgccaac	4800
aucucuggca	acgueegeuu	gggcucuaug	agaaucacag	gacccaaaac	cugcaugaac	4860
accuggcagg	ggaccuuucc	uaucaauugu	uauacagaag	gccagugcuu	gccgaaaccc	4920
gcguuaaacu	ucaagaccgc	caucuggaga	guggcggccu	cagaguacgc	ggaagugacg	4980
cagcacggau	cauaugccua	uauaacaggg	cugaccacug	acaacuuaaa	agucccuugc	5040
caacuccccu	cuccagaguu	uuucucuugg	guggacggag	uacaaaucca	uagguccgcc	5100
cccacaccaa	agccguuuuu	ccgggaugag	gucucguuca	gcguugggcu	caauucauuu	5160
gucgucgggu	cucageuuce	cugugacccu	gagcccgaca	cugagguagu	gauguccaug	5220
cuaacagacc	caucccauau	cacggcggag	gcugcagcgc	ggcguuuagc	gcggggguca	5280
ccccaucug	aggcaagcuc	cucagegage	cagcugucgg	cgccaucgcu	gcgagccacc	5340
ugcaccaccc	acgguaggac	cuaugaugug	gacauggugg	augecaaceu	guucaugggg	5400
ggcggcguga	uucggauaga	gucugagucc	aaaguggucg	uucuggacuc	ccucgacuca	5460
augaccgagg	aagagggcga	ccuugagccu	ucaguaccau	cggaguauau	geueeecagg	5520
aagagguucc	caccggccuu	accggcuugg	gegeggeeug	auuacaaccc	accgcuugug	5580
gaaucgugga	agaggccaga	uuaccaacca	eccacuguug	cgggcugugc	ucucccccc	5640
cccaaaaaga	ccccgacgec	uccuccaagg	agaegeegga	cagugggucu	gagcgagagc	5700
accauaggag	augeceucea	acageuggee	aucaaguccu	uuggccagcc	cccccaagc	5760
ggcgauucag	geeuuueeae	gggggcggac	geegeegaeu	ccggcgaucg	gacacccccu	5820
gacgaguugg	cucuuucgga	gacagguucu	accuccucca	ugecececcu	cgagggggag	5880
ccuggggacc	cagaccugga	gccugagcag	guagagcuuc	aaccuccucc	ccaggggggg	5940
gaggcagcuc	ceggeuegga	cucggggucc	uggucuacuu	geueegagga	ggaugacucc	6000
gucgugugcu	gcuccauguc	auauuccugg	accggggcuc	uaauaacucc	uuguageeee	6060
gaagaggaaa	aguugccaau	uaacuccuug	agcaacucgc	uguugcgaua	ccauaacaag	6120
guauacugua	cuacaucaaa	gagugecuca	cuaagggcua	aaaagguaac	uuuugauagg	6180
augcaagugc	ucgacgccua	uuaugauuca	gucuuaaagg	acaucaagcu	ageggeeuce	6240
aaggucagcg	caaggcuccu	caccuuagag	gaggcgugcc	aauugacccc	accccacucu	6300
gcaagaucca	aguauggguu	uggggcuaag	gagguccgca	gcuuguccgg	gagggccguc	6360
aaccacauca	aguccgugug	gaaggaccuc	uuggaagacu	cacaaacacc	aauuccuaca	6420
accaucaugg	ccaaaaauga	gguguucugc	guggaccccg	ccaagggggg	uaaaaaacca	6480
gcucgccuua	ucguuuaccc	ugaccucggc	gucagggucu	gcgagaagau	ggcccuuuau	6540
gaugucacac	aaaagcuucc	ucaggcggug	augggggcuu	cuuauggcuu	ccaguacucc	6600
cccgcucage	ggguggaguu	ucucuugaag	gcaugggcgg	aaaagagaga	cccuaugggu	6660

```
uuuucguaug auacccgaug cuuugacuca accgucacug agagagacau caggacugag 6720
gaguccauau accaggocug cuccuuacce gaggaggocc gaacugccau acacucgcug 6780
acugagagac ucuauguggg agggcccaug uucaacagca agggccaguc cugcggguac 6840
aggeguugee gegeeagegg ggugeuuace acuaguaugg ggaacaccau cacaugeuau 6900
guaaaagccc uagcggcuug caaggcugcg gggauaauug cgcccacgau gcugguaugc 6960
ggogacgacu ugquogucau cucagaaago caggggacug aggaggacga goggaaccug 7020
agageeuuca eggaggeuau gaccagguau ueugeeeeue euggugaeee eeccagaeeg 7080
gaauaugacc uggagcuaau aacaucuugu uccucaaacg ugucuguggc acuuggccca 7140
cagggeegee geagauacua ceugaeeaga gaeeceaeea cuucaauuge eegggeugee 7200
ugggaaacag uuagacacuc cccugucaau ucauggcugg gaaacaucau ccaguacgcu 7260
ccaaccauau ggguucgcau gguccugaug acacacuucu ucuccauucu cauggcccag 7320
gacacceuag accagaaccu uaacuuugaa auguacggau egguguacue egugaguecu 7380
cuggaccucc cagccauaau ugaaagguua cacgggcuug acgccuucuc ucugcacaca 7440
uacacucece acgaacugae geggguggeu ucageecuca gaaaacuugg ggegeeacee 7500
cucagagegu ggaagagueg ggegegugea guuagggegu eccucaucue eegugggggg 7560
agggeggeeg uuugeggueg guaecucuuc aacugggegg ugaagaccaa gcucaaacuc 7620
acuccuunge eggaggeacg cenecuggan unguccagun ggunnacegn eggegeegge 7680
gggggcgaca uuuaucacag cgugucgcgu gcccgacccc gccuauuacu ccuuagccua 7740
cuccuacuuu cuguaggggu aggccucuuc cuacuccccg cucgauagag cggcacacau 7800
uuuuuuuuu cuuuuuuuu uuuuucccuc uuucuucccu ucucaucuua uucuacuuuc 7920
uuucuuggug geuceaucuu ageeeuague aeggeuageu gugaaaggue egugageege 7980
augacugcag agagugcegu aacuggucuc ucugcagauc augu
                                                             8024
```

```
<210> 3
```

<211> 9678

<212> DNA

<213> Hepatitis C virus

<220>

<221> CDS

<222> (341)..(9442)

10

<400> 3

acctgccct aatagggcg acactccgcc atgaatcact cccctgtgag gaactactgt 60
cttcacgcag aaagcgccta gccatggcgt tagtatgagt gtcgtacagc ctccaggccc 120
ccccctcccg ggagagccat agtggtctgc ggaaccggtg agtacaccgg aattgccggg 180
aagactgggt cctttcttgg ataaacccac tctatgcccg gccatttggg cgtgcccccg 240
caagactgct agccgagtag cgttgggttg cgaaaggcct tgtggtactg cctgataggg 300
cgcttgcgag tgccccggga ggtctcgtag accgtgcacc atg agc aca aat cct 355
Met Ser Thr Asn Pro

aaa cct caa aga aaa acc aaa aga aac acc aac cgt cgc cca gaa gac 403 Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn Arg Arg Pro Glu Asp

15

1

2.0

gtt aag tte eeg gge gge eag ate gtt gge gga gta tae ttg ttg 451
Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly Gly Val Tyr Leu Leu
25 30 35

ccg cgc agg ggc ccc agg ttg ggt gtg cgc acg aca agg aaa act tcg 499

Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Thr Thr Arg Lys Thr Ser

40 45 50

gag cgg tcc cag cca cgt ggg aga cgc cag ccc atc ccc aaa gat cgg 547

Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro Ile Pro Lys Asp Arg

55 60 65

cgc	tcc	act	ggc	aag	gee	tgg	gga	aaa	cca	ggt	cgc	ccc	tgg	ccc	cta	595
Arg	Ser	Thr	Gly	Lys	Ala	Trp	Gly	ГÀВ	Pro	Gly	Arg	Pro	Trp	Pro	Leu	
70				7	5				80				8 5			
tat	ggg	aat	gag	gga	¢tc	ggc	tgg	gca	gga	tgg	ctc	ctg	tcc	ccc	cga	643
Tyr	Glγ	Asn	Glu	Gly	Leu	Gly	Trp	Ala	Gly	Trp	Leu	Leu	Ser	Pro	Arg	
			9	0				95				100				
	-															
ggc	tct	cgc	acc	tcc	tgg	ggc	ccc	act	gac	ccc	cgg	cat	agg	tcg	cgc	691
Gly	Ser	Arg	Pro	Ser	Trp	Gly	Pro	Thr	Asp	Pro	Arg	His	Arg	Ser	Arg	
		1	05				110				11	5				
aac	gtg	ggt	aaa	gtc	atc	gac	acc	cta	acg	tgt	ggc	ttt	gcc	gac	ctc	739
Asn	Val	Gly	Lys	Val	Ile	Asp	Thr	Leu	Thr	Cys	Glγ	Phe	Ala	Asp	Leu	
	1	.20				125				13	0					
atg	aaa	tac	atc	ccc	gtc	gta	ggc	gcc	ccg	ctt	agt	ggc	gcc	gcc	aga	787
Met	Gly	Tyr	Ile	Pro	Val	Val	Gly	Ala	Pro	Leu	Ser	Gly	Ala	Ala	Arg	
	135				140	)			1	4 5						
gct	gtc	gcg	cac	ggc	gtg	aga	gtc	ctg	gag	gaç	999	gtt	aat	tat	gca	835
Ala	Val	Ala	His	Gly	Val	Arg	Val	Leu	Glu	Asp	Gly	Val	Asn	Tyr	Ala	
150				15	5				L 6 0				165			
aca	ggg	aac	cta	ccc	ggt	ttc	ccc	ttt	tct	atc	ttc	ttg	ctg	gcc	ctg	883
Thr	Gly	Asn	Leu	Pro	Gly	Phe	Pro	Phe	Ser	Ile	Phe	Leu	Leu	Ala	Leu	
			17	0			=	175				180				
ttg	tcc	tge	atc	acc	gtt	ccg	gtc	tet	gct	gcc	cag	gtg	aag	aat	acc	931
Leu	Ser	Сув	Ile	Thr	Val	Pro	Val	Ser	Ala	Ala	Gln	Val	Lys	Asn	Thr	
		1.8	35				190				199	õ				

agt	agc	agc	tac	atg	gtg	acc	aat	gac	tgc	tcc	aat	gac	agc	atc	act	979
Ser	Ser	Ser	Tyr	Met	Val	Thr	Asn	Asp	Сув	Ser	Asn	Asp	Ser	Ile	Thr	
	2	00				205				21	0					
tgg	cag	ctc	gag	gat	gcg	gtt	ctc	cac	gtc	ccc	aaa	tgc	gtc	ccg	tgc	1027
Trp	Gln	Leu	Glu	Ala	Ala	Val	Leu	His	Val	Pro	Gly	Сув	Val	Pro	Сла	
	215				220	)			2	25						
gag	aga	gtg	999	aat	acg	tca	cgg	tgt	tgg	gtg	cca	gtc	tcg	cca	aac	1075
Glu	Arg	Val	Gly	Asn	Thr	Ser	Arg	Cys	Trp	Val	Pro	Val	ser	Pro	Asn	
230				23	5			:	240				245			
atg	get	gtg	cgg	cag	ccc	ggt	gcc	ctc	acg	cag	ggt	ctg	cgg	acg	cac	1123
Met	Ala	Val	Arg	Gln	Pro	Gly	Ala	Leu	Thr	Gln	Gly	Leu	Arg	Thr	His	
			25	0			:	255				260				
atc	gat	atg	gtt	gtg	atg	tee	gcc	acc	ttc	tgc	tet	gct	ctc	tac	gtg	1171
Ile	Asp	Met	Val	Val	Met	Ser	Ala	Thr	Phe	Сув	Ser	Ala	Leu	Tyr	Val	
		2	65				270				27	5				
999	gac	ata	tgt	ggc	999	gtg	atg	ctc	gcg	gcc	cag	gtg	ttc	atc	gtc	1219
Gly	Asp	Leu	Cys	Glγ	Gly	Val	Met	Leu	Ala	Ala	Gln	Val	Phe	Ile	Val	
	2	80				285				29	0					
tcg	ccg	cag	tac	cac	tgg	ttt	gtg	caa	gaa	tgc	aat	tgc	tcc	atc	tac	1267
Ser	Pro	Gln	Tyr	His	Trp	Phe	Val	Gln	Glu	Сув	Asn	Сув	Ser	Ile	Tyr	
	295				300	)			3	05						
cct	ggc	acc	atc	act	gga	cac	cgc	atg	gca	tgg	gac	atg	atg	atg	aac	1315
Pro	Gly	Thr	Ile	Thr	Gly	His	Arg	Met	Ala	Trp	Asp	Met	Met	Met	Asn	
310				31	5			3	320				325			

tgg	tcg	ccc	acg	gcc	acc	atg	atc	ctg	gcg	tac	gtg	atg	cgc	gtc	ccc	1363
Trp	Ser	Pro	Thr	Ala	Thr	Met	Ile	Leu	Ala	Tyr	Va 1	Met	Arg	Val	Pro	
			3 3	0				335				340				
gag	gtc	atc	ata	gac	atc	gtt	agc	<u>a</u> aa	gct	cac	tgg	ggc	gtc	atg	ttc	1411
Glu	Val	Ile	Ile	Asp	Ile	Val	Ser	Gly	Ala	His	Trp	Gly	Val	Met	Phe	
		3	45				350				35	5				
ggc	ttg	gcc	tac	ttc	tet	atg	cag	gga	gcg	tgg	gcg	aag	gtc	att	gtc	1459
Gly	Leu	A1a	Tyr	Phe	Ser	Met	Gln	Gly	Ala	Trp	Ala	Lys	Va1	Ile	Val	
	3	360				365				37	0					
atc	ctt	ctg	ctg	gcc	gct	ggg	gtg	gac	gcg	ggc	acc	acc	acc	gtt	gga	1507
Ile	Leu	Leu	Leu	Ala	Ala	Gly	Val	Asp	Ala	Gly	Thr	Thr	Thr	Val	Gly	
	375				380	)			3	85						
ggc	gct	gtt	gca	cgt	tcc	acc	aac	gtg	att	gcc	ggc	gtg	ttc	agc	cat	1555
Gly	Ala	Val	Ala	Arg	Ser	Thr	Asn	Val	Ile	Ala	Gly	Val	Phe	Ser	His	
390				3 9	5			4	00				405			
ggc	cct	cag	cag	aac	att	cag	ctc	att	aac	acc	aac	ggc	agt	tgg	cac	1603
Gly	Pro	Gln	Gln	Asn	Ile	Gln	Leu	Ile	Asn	Thr	Asn	Gly	Ser	Trp	His	
			41	0			•	115				420				
atc	aac	cgt	act	gcc	ttg	aat	tgc	aat	gac	tcc	ttg	aac	acc	ggc	ttt	1651
Ile	Asn	Arg	Thr	Ala	Leu	Asn	Cys	Asn	Asp	Ser	Leu	Asn	Thr	Gly	Phe	
		4	25				430				43	5				
ctc	gcg	gcc	ttg	ttc	tac	acc	aac	cgc	ttt	aac	tcg	tca	333	tgt	cca	1699
Leu	Ala	Ala	Leu	Phe	Tyr	Thr	Asn	Arg	Phe	Asn	Ser	Ser	Gly	Cys	Pro	
	4	40				445				45	0					

999	cgc	ctg	tee	gcc	tgc	cgc	aac	atc	gag	gct	ttc	cgg	ata	999	tgg	1747
Gly	Arg	Leu	Ser	Ala	Сув	Arg	Asn	Ile	Glu	Ala	Phe	Arg	Ile	Gly	Trp	
	455				460	)			4	65						
ggc	acc	cta	cag	tac	gag	gat	aat	gtc	acc	aat	cca	gag	gat	atg	agg	1795
Gly	Thr	Leu	Gln	Tyr	Glu	Asp	Asn	Val	Thr	Asn	Pro	Glu	Asp	Met	Arg	
470				47	5			4	180				485			
ccg	tac	tgc	tgg	cac	tac	ccc	cca	aag	ccg	tgt	ggc	gta	gtc	ccc	gcg	1843
Pro	Tyr	Cys	Trp	His	тут	Pro	Pro	Lys	Pro	Сув	Gly	Val	Val	Pro	Ala	
			49	0				495		٠		500				
agg	tct	gtg	tgt	ggc	cca	gtg	tac	tgt	ttc	acc	ccc	agc	ccg	gta	gta	1891
Arg	Ser	Val	Cys	Gly	Pro	Val	Tyr	Cys	Phe	Thr	Pro	Ser	Pro	Val	Val	
		5	05				510				51	5				
gtg	ggc	acg	acc	gac	aga	cgt	gga	gtg	ccc	acc	tac	aca	tgg	gga	gag	1939
Val	Gly	Thr	Thr	Asp	Arg	Arg	Gly	Val	Pro	Thr	Tyr	Thr	Trp	Gly	Glu	
	5	20				525				53	0					
aat	gag	aca	gat	gtc	ttc	cta	ctg	aac	agc	acc	cga	ccg	ccg	cag	ggc	1987
Asn	Glu	Thr	Asp	Val	Phe	Leu	Leu	Asn	Ser	Thr	Arg	Pro	Pro	Gln	Gly	
	535				540	}			5	4 5						
tca	tgg	ttc	ggc	tgc	acg	tgg	atg	aac	tee	act	ggt	ttc	acc	aag	act	2035
Ser	Trp	Phe	Gly	Cys	Thr	Trp	Met	Asn	Ser	Thr	Gly	Phe	Thr	Lys	Thr	
550				5 5	5			į	560				565			
tgt	ggc	gcg	cca	cct	tgc	cgc	acc	aga	gct	gac	ttc	aac	gcc	agc	acg	2083
Сув	Gly	Ala	Pro	Pro	Cys	Arg	Thr	Arg	Ala	Asp	Phe	Asn	Ala	ser	Thr	
			57	0				575				580				

gac	ccg	ccg	cyc	CCC	acg	gac	cgt	CCC	agg	aag	cac	CCL	gat	gcc	act	2133
Asp	Leu	Leu	Cys	Pro	Thr	Asp	Cys	Phe	Arg	Lys	His	Pro	Asp	Ala	Thr	
		5	85				590				59	5				
tat	att	aag	tgt	ggt	tct	ggg	ccc	tgg	ctc	aca	cca	aag	tgc	ctg	gtc	2179
Tyr	lle	Lys	Сув	Gly	Ser	Gly	Pro	Trp	Leu	Thr	Pro	Lys	Сув	Leu	Va1	
	(	500				605				61	LO					
cac	tac	cct	tac	aga	ctc	tgg	cat	tac	ccc	tgc	aca	gtc	aat	ttt	acc	2227
His	Tyr	Pro	Tyr	Arg	Leu	Trp	His	Tyr	Pro	Сув	Thr	Val	Asn	Phe	Thr	
	615				620	)			6	25						
atc	ttc	aag	ata	aga	atg	tat	gta	999	999	gtt	gag	cac	agg	ctc	acg	2275
Ile	Phe	Гуз	Ile	Arg	Met	тух	Val	Gly	Gly	Val	Glu	His	Arg	Leu	Thr	
630				63	5			{	5 <b>4</b> ' 0				645			
gcc	gca	tgc	aac	ttc	act	cgt	ggg	gat	cgc	tgc	gac	ttg	gag	gac	agg	2323
Ala	Ala	Cys	Asn	Phe	Thr	Arg	Gly	Asp	Arg	Cys	Asp	Leu	Glu	Asp	Arg	
			65	0			í	555				660				
gac	agg	agt	cag	ctg	tet	cct	ctg	ttg	cac	tct	acc	acg	gaa	tgg	gcc	2371
Asp	Arg	Ser	Gln	Leu	Ser	Pro	Leu	Leu	His	Ser	Thr	Thr	Glu	Trp	Ala	
		6	65				670				67	5				
atc	ctg	ccc	tgc	acc	tac	tca	gac	tta	ccc	gct	ttg	tca	act	ggt	ctt	2419
(le	Leu	Pro	Cys	Thr	Tyr	Ser	Asp	Leu	Pro	Ala	Leu	Ser	Thr	Gly	Leu	
	6	80				685				69	0					
ete	cac	ctt	cac	cag	aac	atc	gtg	gac	gta	caa	tac	atg	tat	ggc	ctc	2467
Leu	His	Leu	His	Gln	Asn	Ile	Val	Asp	Val	Gln	Tyr	Met	туг	Gly	Leu	
	695				700				7	05						

				aca												2513
Ser	Pro	Ala	Ile	Thr	Lys	Tyr	Val	Val	Arg	Trp	Glu	Trp	Val	Val	Leu	
710				71	L 5				720				725	;		
tta	ttc	ctg	ctc	tta	gcg	gac	gcc	aga	gtc	tgc	gcc	tgc	ttg	tgg	atg	2563
Leu	Phe	Leu	Leu	Leu	Ala	Asp	Ala	Arg	Val	Сув	Ala	Cys	Leu	Trp	Met	
			73	30				735				740	)			
ctc	atc	ttg	ttg	ggc	cag	gcc	gaa	gca	gca	ttg	gag	aag	ttg	gtc	gtc	2611
Leu	Ile	Leu	Leu	Gly	Gln	Ala	Glu	Ala	Ala	Leu	Glu	Lys	Leu	Val	Val	
		7	45				750				75	5				
ttg	cac	gat	gcg	agt	gcg	gct	aac	tgc	cat	ggc	ctc	cta	tat	ttt	gcc	2659
Leu	His	Ala	Ala	Ser	Ala	Ala	Asn	Сув	His	Gly	Leu	Leu	Tyr	Phe	Ala	
	•	760				765				77	0					
atc	ttc	ttc	gtg	gca	gct	tgg	cac	atc	agg	ggt	cgg	gtg	gtc	ccc	ttg	2707
Ile	Phe	Phe	Val	Ala	Ala	Trp	His	Ile	Arg	Gly	Arg	Val	Val	Pro	Leu	
	775				780	)			7	85						
acc	acc	tat	tgc	ctc	act	ggc	cta	tgg	ccc	ttc	tgc	cta	ctg	ctc	atg	2755
			Cys													
790				79		-			300		-		805			
gca	ctg	ccc	egg	cag	gct	tat	gcc	tat	gac	qca	cct	ata	cac	gga	cag	2803
			Arg													
			81			•		315				820		1		
				-			`					020				
ata	gac	a t.a	ggt	ttσ	tta	ata	tta	atc	arc	cha	tta	ara	ctc	acc	aaa	2851
			Gly													2011
_ •	1		25				830	110	****	ar Çi Çi	835		Heu	TIIL	r I U	
		٠,									035	,				

aaa	tat	aag	acc	ctc	ctc	ggc	cag	tgt	ctg	tgg	tgg	ttg	tgc	tat	ctc	2899
G1y	Tyr	Lys	Thr	Leu	Leu	Gly	Gln	Сув	Leu	Trp	Trp	Leu	Cys	Tyr	Leu	
	8	340				845				85	0					
ctg	acc	ctg	ggg	gaa	gcc	atg	att	cag	gag	tgg	gta	cca	ccc	atg	cag	2947
Leu	Thr	Leu	Gly	Glu	Ala	Met	Ile	Gln	Glu	Trp	۷al	Pro	Pro	Met	Gln	
	855				860	)			8	65						
gtg	cgc	ggc	ggc	aga	gat	ggc	atc	gcg	tgg	gcc	gtc	act	ata	ttc	tgc	2995
Val	Arg	Gly	Gly	Arg	Asp	Gly	ĭle	Ala	Trp	Ala	Val	Thr	Ile	Phe	Сув	
870				87	5			8	880				885			
ccg	ggt	gtg	gtg	ttt	gac	att	acc	aaa	tgg	ctt	ttg	gcg	ttg	ctt	ggg	3043
Pro	Gly	Val	Val	Phe	Asp	Ile	Thr	гав	Trp	Leu	Leu	Ala	Leu	Leu	Gly	
			89	0			;	895				900				
cct	gct	tac	ctc	tta	agg	gcc	gct	ttg	aca	cat	gtg	ccg	tac	ttc	gtc	3091
Pro	Ala	Tyr	Leu	Leu	Arg	Ala	Ala	Leu	Thr	His	Val	Pro	Tyr	Phe	Val	
		91	05				910				91	5				
aga	gct	cac	gct	ctg	ata	agg	gta	tgc	gct	ttg	gtg	aag	cag	ctc	gcg	3139
Arg	Ala	His	Ala	Leu	Ile		Val	Cys	Ala	Leu	Val	Lys	Gln	Leu	Ala	
	9	20				925				93	0					
				gtt		• "			•	-						3187
Gly	Gly	Arg	Tyr	Val	Gln	Val	Ala	Leu	Leu	Ala	Leu	Gly	Arg	Trp	Thr	
	935				940	•			9	45						
				tat						_	_	-		_	_	3235
-	Thr	Tyr	Ile	Tyr	-	His	Leu			Met	Ser	Asp	-	Ala	Ala	
950				95	5			9	960				965			

agc	ggc	ctg	cgc	gac	tta	gcg	gtc	gcc	gtg	gaa	aac	atc	atc	ttc	agt	3283
Ser	Gly	Leu	Arg	Asp	Leu	Ala	Val	Ala	Val	Glu	Pro	Ile	Ile	Phe	Ser	
			97	0				975				980				
ccg	atg	gag	aag	aag	gtc	atc	gtc	tgg	gga	gcg	gag	acg	gct	gca	tgt	3331
Pro	Met	Glu	Lys	Lys	Val	Ile	Val	Trp	Gly	Ala	Glu	Thr	Ala	Ala	Сув	
		9	85				990				99	5				
ggg	gac	att	cta	cat	gga	ctt	ccc	gtg	tcc	gcc	cga	ctc	ggc	cag	gag	3379
Gly	Asp	Ile	Leu	His	Gly	Leu	Pro	Val	Ser	Ala	Arg	Leu	Gly	Gln	Glu	
	10	00				1005				10	10					
atc	ctc	ctc	ggc	cca	gct	gat	ggc	tac	acc	tcc	aag	333	tgg	aag	ctc	3427
Ile	Leu	Leu	Gly	Pro	Ala	Asp	Gly	Tyr	Thr	ser	Lys	Gly	Trp	Lys	Leu	
1	015				102	0			10	25						
ctt	gct	ccc	atc	act	gct	tat	gcc	cag	caa	aca	cga	ggc	ctc	ctg	ggc	3475
Leu	Ala	Pro	Ile	Thr	Ala	Tyr	Ala	Gln	Gln	Thr	Arg	Gly	Leu	Leu	Gly	
1030	)			103	5			1	040				104	ā		
gcc	ata	gtg	gtg	agt	atg	acg	ggg	cgt	gac	agg	aca	gaa	cag	gcc	ggg	3523
Ala	Ile	Val	Val	Ser	Met	Thr	Gly	Arg	Asp	Arg	Thr	Glu	Gln	Ala	Gly	
			105	0			1	055				1066	)			
gaa	gtc	caa	atc	ctg	tee	aca	gtc	tct	cag	tcc	ttc	ctc	gga	aca	acc	3571
Glu	Val	Gln	Ile	Leu	Ser	Thr	Val	Ser	Gln	Ser	Phe	Leu	Gly	Thr	Thr	
		10	6 5			1	070				107	5				
atc	teg	aaa	gtt	ttg	tgg	act	gtt	tac	cac	gga	gct	ggc	aac	aag	act	3619
Tle	Ser	Gly	Val	Leu	Trp	Thr	Val	Tyr	His	Gly	Ala	Gly	Asn	Lys	Thr	
	10	80				1085				109	9 0					

cta	gcc	ggc	tta	cgg	ggt	ccg	gtc	acg	cag	atg	tac	tcg	agt	gct	gag	3667
Leu	Ala	Gly	Leu	Arg	Gly	Pro	Val	Thr	Gln	Met	Tyr	Ser	Ser	Ala	Glu	
1	095				110	0			11	.05						
999	gac	ttg	gta	ggc	tgg	ccc	agc	ccc	cct	ggg	acc	aag	tet	ttg	gag	3715
Gly	Asp	Leu	Val	Gly	Trp	Pro	Ser	Pro	Pro	Gly	Thr	Lуs	Ser	Leu	Glu	
1110	0			111	L <b>5</b>			1	120				112	5		
ccg	tgc	aag	tgt	gga	gcc	gtc	gac	cta	tat	ctg	gtc	acg	cgg	aac	get	3763
Pro	Cys	Lys	Сув	Gly	Ala	Val	Asp	Leu	Tyr	Leu	Val	Thr	Arg	Asn	Ala	
			113	0			1	135				114	0			
gat	gtc	atc	ccg	get	cgg	aga	cgc	999	gac	aag	cgg	gga	gca	ttg	ctc	3811
Asp	Val	Ile	Pro	Ala	Arg	Arg	Arg	Gly	Asp	ГАЗ	Arg	Gly	Ala	Leu	Leu	
		11	45			5	1150				115	5				
tee	ccg	aga	ccc	att	teg	acc	ttg	aag	999	tee	tcg	<b>a</b> aa	<b>ggg</b>	ccg	gtg	3859
Ser	Pro	Arg	Pro	Ile	Ser	Thr	Leu	Lys	Gly	Ser	Ser	Gly	Gly	Pro	Va1	
	1.7	160				1165				11	70					
ctc	tgc	cct	agg	ggc	cac	gtc	gtt	999	ctc	ttc	ega	gca	gct	gtg	tgc	3907
Leu	Cys	Pro	Arg	Gly	Ris	Va1	Val	Gly	Leu	Phe	Arg	Ala	Ala	Val	Сув	
1	175				118	0			11	85						
tet	cgg	ggc	gtg	gcc	aaa	tcc	atc	gat	ttc	atc	ccc	gtt	gag	aca	ctc	3955
Ser	Arg	Gly	Val	Ala	Lys	Ser	Ile	Asp	Phe	Ile	Pro	Va1	Glu	Thr	Leu	
1190	)			119	5			1	200				120	5		
gac	gtt	gtt	aca	agg	tct	ccc	act	ttc	agt	gac	aac	agc	acg	cca	ccg	4003
Asp	Val	Val	Thr	Arg	Ser	Pro	Thr	Phe	ser	Asp	Asn	Ser	Thr	Pro	Pro	
	1210						1	215				1220	)			

get	gtg	ccc	cag	acc	tat	cag	gtc	aaa	tac	ttg	cat	gct	cca	act	ggc	4051
Ala	Val	Pro	Gln	Thr	Tyr	Gln	Val	Gly	Tyr	Leu	His	Ala	Pro	Thr	Gly	
		12	25			1	L230				123	5				
agt	gga	aag	agc	acc	aag	gtc	cct	gtç	gcg	tat	gcc	gcc	cag	999	tac	4099
Ser	Gly	Lys	Ser	Thr	Lys	Val	Pro	Val	Ala	Tyr	Ala	Ala	Gln	Gly	Tyr	
	1:	240				1245				12	50					
aaa	gta	cta	gtg	ctt	aac	ccc	tcg	gta	gct	gcc	acc	ctg	999	ttt	ggg	4147
Lys	Val	Leu	Val	Leu	Asn	Pro	Ser	Val	Ala	Ala	Thr	Leu	Gly	Phe	Gly	
1	255				126	0			12	265						
gcg	tac	cta	tcc	aag	gca	cat	ggc	atc	aat	ccc	aac	att	agg	act	gga	4195
Ala	Tyr	Leu	Ser	Lys	Ala	His	Gly	Ile	Asn	Pro	Asn	Ile	Arg	Thr	Gly	
1276				127	5			1	280				128	5		
gtc	agg	acc	gtg	atg	acc	<b>9</b> 99	gag	gee	atc	acg	tac	tee	aca	tat	ggc	4243
Val	Arg	Thr	Val	Met	Thr	Gly	Glu	Ala	Ile	Thr	Tyr	Ser	Thr	Tyr	Gly	
			129	0			1	295				130	0			
aaa	ttt	ctc	gcc	gat	999	ggc	tgc	gct	agc	ggc	gcc	tat	gac	atc	atc	4291
Lys	Phe	Leu	Ala	Asp	Gly	Gly	Cys	Ala	ser	Gly	Ala	Tyr	Asp	Ile	Ile	
		13	05			1	1310				131	5				
ata	tgc	gat	gaa	tgc	cac	get	gtg	gat	gct	acc	tee	att	ctc	ggc	atc	4339
Ile	Cys	Asp	Glu	Cys	His	Ala	Val	Asp	Ala	Thr	Ser	Ile	Leu	Gly	Ile	
	13	320				1325				13	3 0					
gga	acg	gtc	ctt	gat	caa	gca	gag	aca	gcc	<b>3</b> 33	gtc	aga	cta	act	gtg	4387
Gly	Thr	Val	Leu	Asp	Gln	Ala	Glu	Thr	Ala	Gly	Val	Arg	Leu	Thr	Val	
1	335				1340	)			13	45						

ctg	gct	acg	gcc	aca	ccc	ccc	ggg	tca	gtg	aca	acc	ccc	cat	aca	gat	4435
Leu	Ala	Thr	Ala	Thr	Pro	Pro	Gly	Ser	Val	Thr	Thr	Pro	His	Pro	Asp	
1350	)			135	5 5			1	360				136	j j		
ata	gaa	gag	gta	ggc	ctc	999	cgg	gag	ggt	gag	atc	ccc	ttc	tat	<b>9</b> 99	4483
Ile	Glu	Glu	Val	Gly	Leu	Gly	Arg	Glu	Gly	Glu	Ile	Pro	Phe	Tyr	Gly	
			137	70			1	375				138	0			
agg	gcg	att	ccc	cta	tcc	tgc	atc	aag	gga	999	aga	cac	ctg	att	ttc	4531
Arg	Ala	Ile	Pro	Leu	Ser	Сув	Ile	ьув	Gly	Gly	Arg	His	Leu	Ile	Phe	
		13	85				1390				139	5				
tgc	cac	tca	aag	aaa	aag	tgŧ	gac	gag	ctc	gcg	gcg	gcc	ctt	cgg	ggc	4579
Cys	His	Ser	ГХS	Lys	гуs	Cys	Asp	Glu	Leu	Ala	Ala	Ala	Leu	Arg	Gly	
	14	100				1405				14	10					
atg	ggc	ttg	aat	gee	gtg	gca	tac	tat	aga	999	ttg	gac	gtc	tcc	ata	4627
Met	Gly	Leu	Asn	Ala	Va1	Ala	Tyr	Tyr	Arg	Gly	Leu	Asp	Val	Ser	Ile	
1	415				142	0			14	25						
ata	cca	gct	cag	gga	gat	gtg	gtg	gtc	gtc	gcc	acc	gac	gcc	ata	atg	4675
Ile	Pro	Ala	Gln	Gly	Asp	Val	Val	Val	Val	Ala	Thr	Asp	Ala	Leu	Met	
1430	)			143	5			1	440				1445	5		
acg	aaa	tac	act	gga	gac	ttt	gac	tcc	gtg	atc	gac	tgc	aat	gta	gcg	4723
Thr	Gly	Tyr	Thr	Gly	Asp	Phe	Asp	Ser	Val	Ile	Asp	Cys	Asn	Val	Ala	
			145	0			1	455				1460	)			
gtc	acc	caa	gct	gtc	gac	ttc	agc	ctg	gac	ccc	acc	ttc	act	ata	acc	4771
Val	Thr	Gln	Ala	Val	Asp	Phe	Ser	Leu	Asp	Pro	Thr	Phe	Thr	Ile	Thr	
		14	65			1	470				147	5				

aca	cag	act	gtc	cca	caa	gac	gct	gtc	tca	cgc	agt	cag	cgc	cgc	999	4819
Thr	Gln	Thr	Va1	Pro	Gln	Asp	Ala	Val	Ser	Arg	Ser	Gln	Arg	Arg	Gly	
	14	480				1485				14	90					
cgc	aca	ggt	aga	gga	aga	cag	ggc	act	tat	agg	tat	gtt	tee	act	ggt	4867
Arg	Thr	Gly	Arg	Gly	Arg	Gln	Gly	Thr	Туг	Arg	Tyr	Val	Ser	Thr	Gly	
1	495				1.50	0			15	05						
gaa	cga	gcc	tca	gga	atg	ttt	gac	agt	gta	gtg	ctt	tgt	gag	tgc	tac	4915
Glu	Arg	Ala	Ser	Gly	Met	Phe	Asp	Ser	Val	Val	Leu	Cys	Glu	Cys	Tyr	
151	0			151	. 5			1	520				152	5		
gac	gca	999	gct	gcg	tgg	tac	gat	ctc	aca	cca	gcg	gag	acc	acc	gtc	4963
Asp	Ala	Gly	Ala	Ala	Trp	Tyr	Asp	Leu	Thr	Pro	Ala	Glu	Thr	Thr	Val	
			153	0			1	535				154	)			
agg	ctt	aga	gcg	tat	ttc	aac	acg	ccc	ggc	cta	aaa	gtg	tgt	caa	gac	5011
Arg	Leu	Arg	Ala	туг	Phe	Asn	Thr	Pro	Gly	Leu	Pro	Val	Сув	Gln	Asp	
		15	45			1	550				155	5				
cat	ctt	gaa	ttt	tgg	gag	gca	gtt	ttc	acc	ggc	ctc	aca	cac	ata	gac	5059
Ris	Leu	Glu	Phe	Trp	Glu	Ala	Val	Phe	Thr	Gly	Leu	Thr	His	Ile	Asp	
	15	560				1565				15	70					
gcc	cac	ttc	ctc	tcc	caa	aça	aag	caa	g¢g	999	gag	aac	ttc	gcg	tac	5107
Ala	His	Phe	Leu	Ser	Gln	Thr	гув	Gln	Ala	Gly	Glu	Asn	Phe	Ala	Tyr	
1	575				1580	)			15	85						
cta	gta	gcc	tac	caa	gat	acg	gtg	tgc	gcc	aga	gcc	aag	gcc	cct	ccc	5155
Leu	Val	Ala	Tyr	Gln	Ala	Thr	∀al	Сув	Ala	Arg	Ala	Lys	Ala	Pro	Pro	
1590	}			159	5			1	600				1609	<del>.</del>		

ccg	tee	tgg	gac	gcc	atg	tgg	aag	tgc	ctg	gcc	cga	ctc	aag	cct	acg	5203
Pro	Ser	Trp	Asp	Ala	Met	Trp	Lys	СЛа	Leu	Ala	Arg	Leu	Lys	Pro	Thr	
			161	LO			1	615				162				
ctt	gcg	ggc	ccc	aca	cct	ctc	ctg	tac	cgt	ttg	ggc	cct	att	acc	aat	5251
Leu	Ala	Gly	Pro	Thr	Pro	Leu	Leu	Tyr	Arg	Leu	Gly	Pro	Ile	Thr	Asn	
		16	25				1630				163	5				
gag	gtc	acc	ctc	aca	cac	cct	999	acg	aag	tac	atc	gcc	aca	tgc	atg	5299
Glu	Val	Thr	Leu	Thr	His	Pro	Gly	Thr	Lys	Tyr	Ile	Ala	Thr	Cys	Met	
	1	640				1645	į.			16	50					
				gag				-			_					5347
		Asp	Leu	Glu			Thr	Ser		_	Val	Leu	Ala	Gly	Gly	
1	655				166	0			16	65						
				gtc -												5395
		Ala	Ala	Val		Ala	Tyr			Ala	Thr	Gly	_		Ser	
1670 16					75			1	680				168	ō		
				ttg						_	-	_		_	_	5443
ile	116	Gly	_	Leu	HIS	Val			Arg	Val	vai			Pro	Asp	
			169	, 0			1	695				1700	,			
																F 4 0 1
				tat												5491
пув	GIU			Туг	Gru			Asp	Giu	мес			сув	AIa	ser	
		17	05			_	1710				171	5				
agg	aca	act	ete	atc	gaa	aea	aaa	nan	caa	ata	acc	asa	at~	tta	224	5539
				Ile												2229
ur A			шeи	116		1725	GTÅ	2711	vrâ	173		aru	MEL	пеп	чүр	
1720						1123				11.						

tcc	aag	atc	caa	ggc	ttg	ctg	cag	cag	gcc	tct	aag	cag	gcc	cag	gac	5587		
ser	Lys	Ile	Gln	Gly	Leu	Leu	Gln	Gln	Ala	Ser	Ъуs	Gln	Ala	Gln	Asp			
1	735				174	0			1745									
ata	caa	ccc	gct	atg	cag	gct	tca	tgg	ccc	aaa	gtg	gaa	caa	ttt	tgg	5635		
Ile	Gln	Pro	Ala	Met	Gln	Ala	Ser	Trp	Pro	Lys	Val	G1u	Gln	Phe	Trp			
175	1750 175							1	760				176	5				
gaa	aga	cac	atg	tgg	aac	ttc	att	agc	ggc	atc	caa	tac	ctc	gca	gga	5683		
Ala	Arg	His	Met	Trp	Asn	Phe	Ile	Ser	Gly	Ile	Gln	Tyr	Leu	Ala	Gly			
			177	0			1	775				1780	•					
ttg	tca	aca	ctg	cca	999	aac	ccc	gcg	gtg	gct	tcc	atg	atg	gca	ttc	5731		
Leu	Ser	Thr	Leu	Pro	Gly	Asn	Pro	Ala	Val	Ala	ser	Met	Met	Ala	Phe			
1785						-	1790				179	5						
agt	gcc	gcc	ata	acc	agt	ccg	ttg	teg	acc	agt	acc	acc	atc	ctt	ctc	5779		
Ser	Ala	Ala	Leu	Thr	Ser	Pro	Leu	Ser	Thr	Ser	Thr	Thr	Ile	Leu	Leu			
1800					1805				18	10								
aac	atc	atg	gga	ggc	tgg	tta	gcg	tcc	cag	atc	gca	cca	ccc	gcg	999	5827		
		Met	Gly	Gly	Trp	Leu	Ala	Ser	Gln	Ile	Ala	Pro	Pro	Ala	Gly			
1815					182	0			18	325								
														ggc		5875		
		Gly	Phe			Ser	Gly			Gly	Ala	Ala		Gly	Ser			
1830 183					5			1	840				1845					
														ggt		5923		
Ile	Gly	Leu	-	-	Val	Leu		-	Ile	Leu	Ala	-	-	Gly	Ala			
1850							1	855				1860	1860					

ggc	att	teg	aaa	gcc	ctc	gtc	gca	ttc	aag	atc	atg	tct	ggc	gag	aag	5971		
Gly	Ile	Ser	Gly	Ala	Leu	Val	Ala	Phe	Lys	Ile	Met	Ser	Gly	Glu	Lys			
		18	65			-	1870				187	5						
CCC	tet	atg	gaa	gat	gtc	atc	aat	cta	ctg	cct	<u>aaa</u>	atc	ctg	tct	ccg	6019		
Pro	Ser	Met	Glu	Asp	Val	Ile	Asn	Leu	Leu	Pro	Gly	Ile	Leu	Ser	Pro			
	18	880				1885				18	90							
gga	gcc	ctg	gtg	gtg	aaa	gtc	atc	tgc	gcg	gcc	att	ctg	cgc	cgc	cac	6067		
Gly	Ala	Leu	Val	Val	Gly	Val	Ile	Сув	Ala	Ala	Ile	Leu	Arg	Arg	His			
1	895				190	0			19	05								
gtg	gga	ccg	ggg	gag	ggc	gcg	gtc	caa	tgg	atg	aac	agg	ctt	att	gcc	6115		
Val.	Gly	Pro	Gly	Glu	Gly	Ala	Val	Gln	Trp	Met	Asn	Arg	Leu	Ile	Ala			
1910	)			191	.5			1	920				1925					
ttt	gct	tec	aga	gga	aac	cac	gtc	gcc	cct	act	cac	tac	gtg	acg	gag	6163		
Phe	Ala	Ser	Arg	Gly	Asn	His	Val	Ala	Pro	Thr	His	Tyr	Val	Thr	Glu			
			193	0			1	935				1940	)					
tcg	gat	gcg	tcg	cag	cgt	gtg	acc	caa	cta	ctt	ggc	tct	ctt	act	ata	6211		
Ser	Asp	Ala	Ser	Gln	Arg	Val	Thr	Gln	Leu	Leu	Gly	Ser	Leu	Thr	Ile			
		19	4 5			]	950				195	5						
acc	agc	cta	ctc	aga	aga	ctc	cac	aat	tgg	ata	act	gag	gac	tgc	ccc	6259		
Thr	Ser	Leu	Leu	Arg	Arg	Leu	His	Asn	Trp	Ile	Thr	Glu	Asp	Cys	Pro			
	19	960				1965				19'	70							
								_				_			tgc	6307		
		Сув	Ser	Gly		-	Leu	Arg	•		Trp	Asp	Trp	Val	Cys			
1	975				1986	3			7.9	85								

acc	atc	ttg	aca	gac	ttc	aaa	aat	tgg	ctg	acc	tct	aaa	ttg	tte	ccc	6355		
Thr	Ile	Leu	Thr	Asp	Phe	Lys	Asn	Trp	Leu	Thr	Ser	Lys	Leu	Phe	Pro			
1990 1995								2	000	2005								
aag	ctg	ccc	ggc	ctc	ccc	ttc	atc	tet	tgt	caa	aag	999	tac	aag	ggt	6403		
Lys	Leu	Pro	Gly	Leu	Pro	Phe	Ile	Ser	Суз	Gln	Lys	Gly	Tyr	Lys	Gly			
			201	. 0			2	015				2020	)					
gtg	tgg	gcc	ggc	act	ggc	atc	atg	acc	acg	cgc	tgc	cct	tgc	ggc	gcc	6451		
Val	Trp	Ala	Gly	Thr	Gly	Ile	Met	Thr	Thr	Arg	Cys	Pro	Cys	Gly	Ala			
		20	25			2	2030				203	5						
aac	atc	tct	ggc	aat	gtc	cgc	ctg	ggc	tet	atg	agg	atc	aca	ggg	cct	6499		
Asn	Ile	ser	Gly	Asn	Val	Arg	Leu	Gly	Ser	Met	Arg	Ile	Thr	Gly	Pro			
	20	040				2045				20	50							
aaa	acc	tgc	atg	aac	acc	tgg	cag	999	acc	ttt	cct	atc	aat	tgc	tac	6547		
ГХа	Thr	Cys	Met	Asn	Thr	Trp	Gln	Gly	Thr	Phe	Pro	Ile	Asn	Cys	Tyr			
2055					206	0			20	65								
acg	gag	ggc	cag	tgc	gcg	ccg	aaa	ccc	ccc	acg	aac	tac	aag	acc	gcc	6595		
Fhr	Glu	Gly	Gln	Cys	Ala	Pro	Lys	Pro	Pro	Thr	Asn	Tyr	Lys	Thr	Ala			
2070	070 2075							2	080				5					
atc	tgg	agg	gtg	gcg	gcc	tcg	gag	tac	gcg	gag	gtg	acg	cag	cat	aaa	6643		
Ile	тгр	Arg	Val	Ala	Ala	ser	G1u	Tyr	Ala	Glu	Va1	Thr	Gln	His	Gly			
2090							2	095				2100	)					
ccg	tac	tee	tat	gta	aca	gga	ctg	acc	act	gac	aat	ctg	aaa	att	cct	6691		
Ser	Tyr	Ser	туг	Val	Thr	Gly	Leu	Thr	Thr	Asp	Asn	Leu	Lys	Ile	Pro			
	2105					2	110				211	5	5					

tgc	caa	cta	cct	tct	cca	gag	ttt	ttc	tcc	tgg	gtg	gac	ggt	gtg	cag	6739
Cys	Gln	Leu	Pro	Ser	Pro	Glu	Phe	Phe	Ser	Trp	Va1	Asp	Gly	Val	Gln	
	2.	120				2125	•			21	3 0					
	٠															
atc	cat	agg	ttt	gca	ccc	aca	cca	aag	ccg	ttt	ttc	cgg	gat	gag	gtc	6787
Ile	His	Arg	Phe	Ala	Pro	Thr	Pro	ьув	Pro	Phe	Phe	Arg	Asp	Glu	Va1	
2	135	٠			214	0			21	L45						
tcg	ttc	tgc	gtt	ggg	ctt	aat	tcc	tat	gct	gtc	999	tac	cag	ctt	aac	6835
Ser	Phe	Сув	Val	Gly	Leu	Asn	Ser	Tyr	Ala	Val	Glγ	Ser	Gln	Leu	Pro	
215	0			215	55			2	160				216	5		
tgt	gaa	cct	gag	ccc	gac	gca	gac	gta	ttg	agg	tee	atg	cta	aca	gat	6883
Cys	Glu	Pro	Glu	Pro	Asp	Ala	Asp	Val	Leu	Arg	Ser	Met	Leu	Thr	Asp	
	-		217	0			2	175				218	0			
ccg	acc	cac	atc	acg	gcg	gag	act	gcg	gcg	cgg	cgc	ttg	gca	cgg	gga	6931
Pro	Pro	His	Ile	Thr	Ala	Glu	Thr	Ala	Ala	Arg	Arg	Leu	Ala	Arg	Gly	
		21	85			2	2190				219	5				
tca	cct	cca	tct	gag	gcg	agc	tcc	tca	gtg	agc	cag	cta	tca	gca	ccg	6979
Ser	Pro	Pro	Ser	Glu	Ala	Ser	Ser	Ser	Val	Ser	Gln	Leu	Ser	Ala	Pro	
	22	200				2205				22	10					
tog	ctg	cgg	gcc	acc	tgc	acc	acc	cac	agc	aac	acc	tat	gac	gtg	gac	7027
Ser	Leu	Arg	Ala	Thr	Сув	Thr	Thr	His	Ser	Asn	Thr	Tyr	Asp	Val	Asp	
2	215				222	0			22	25						
atg	gtç	gat	gcc	aac	ctg	ctc	atg	gag	ggc	ggt	gtg	gct	cag	aca	gag	7075
Met	Val	Asp	Ala	Asn	Leu	Leu	Met	Glu	Gly	Gly	Va1	Ala	Gln	Thr	Glu	
223(	)			223	5			2	240				2245	ŝ		

cct	gag	tcc	agg	gtg	ccc	gtt	ctg	gac	ttt	ctc	gag	сса	atg	gcc	gag	7123
Pro	Glu	Ser	Arg	Val	Pro	Val	Leu	Asp	Phe	Leu	Glu	Pro	Met	Ala	Glu	
			225	50			2	255				226	0			
gaa	gag	agc	gac	ctt	gag	ccc	tca	ata	cca	tcg	gag	tgc	atg	ata	ccc	7171
G1 u	Glu	Ser	Asp	Leu	Glu	Pro	Ser	Ile	Pro	Ser	Glu	Сув	Met	Leu	Pro	
		22	65			2	2270				227	5				
agg	agc	999	ttt	cca	cgg	gcc	tta	ccg	gct	tgg	gca	cgg	cct	gac	tac	7219
Arg	Ser	Gly	Phe	Pro	Arg	Ala	Leu	Pro	Ala	Trp	Ala	Arg	Pro	Asp	Tyr	
	22	280				2285				22	90					
aac	ccg	ccg	ctc	gtg	gaa	tcg	tgg	agg	agg	cca	gat	tac	caa	ccg	ccc	7267
Asn	Pro	Pro	Leu	Val	Glu	Ser	Trp	Arg	Arg	Pro	Asp	Tyr	Gln	Pro	Pro	
2	295				230	0			23	05						
acc	gtt	gct	ggt	tgt	gct	ctc	ccc	ccc	ccc	aag	aag	gcc	dag	acg	cct	7315
Thr	Val	Ala	Gly	Сув	Ala	Leu	Pro	Pro	Pro	Lys	Lys	Ala	Pro	Thr	Pro	
2310	)			231	. 5			2	320				2325	5		
ccc	cca	agg	aga	cgc	cgg	aca	gtg	ggt	ctg	agc	gag	agc	acc	ata	tca	7363
Pro	Pro	Arg	Arg	Arg	Arg	Thr	Val	Gly	Leu	Ser	Glu	Ser	Thr	Ile	Ser	
			233	0			2	335				2340	)			
gaa	gcc	ctc	cag	caa	ctg	gcc	atc	aag	acc	ttt	ggc	cag	ccc	ccc	tcg	7411
Glu	Ala	Leu	Gln	Gln	Leu	Ala	Ile	Lys	Thr	Phe	Gly	Gln	Pro	Pro	Ser	
		23	45			2	350				235	5				
agc	ggt	gat	gca	ggc	tcg	tcc	acg	<b>3</b> 33	gcg	ggc	gcc	gcc	gaa	tee	ggc	7459
Ser	Gly	Asp	Ala	Gly	Ser	Ser	Thr	Gly	Ala	Gly	Ala	Ala	Glu	ser	Gly	
	23	60				2365				23'	70					

ggt	ccg	acg	tcc	cct	ggt	gag	ccg	gcc	acc	tca	gag	aca	ggt	tee	gcc	7507
Gly	Pro	Thr	Ser	Pro	Gly	Glu	Pro	Ala	Pro	Ser	Glu	Thr	Gly	Ser	Ala	
2	375				238	0			23	85						
tec	tct	atg	cec	ccc	ctc	gag	999	gag	cat	gga	gat	ccg	gac	ctg	gag	7555
Ser	Ser	Met	Pro	Pro	Leu	Glu	Gly	Glu	Pro	Gly	Asp	Pro	Asp	Leu	Glu	
2390	)			239	5			2	400				240	5		
tct	gat	cag	gta	gag	ctt	caa	cct	ccc	ecc	cag	<b>333</b>	<b>a</b> aa	<b>3</b> 33	gta	gct	7603
Ser	Asp	Gln	Val	Glu	Leu	Gln	Pro	Pro	Pro	Gln	Gly	Gly	Gly	Val	Ala	
			241	. 0			2	415				242	О			
aca	ggt	tcg	ggc	tcg	<b>3</b> 33	tct	tgg	tct	act	tgc	taa	gag	gag	gac	gat	7651
Pro	Gly	Ser	Gly	Ser	Gly	Ser	Trp	Ser	Thr	Сув	Ser	Glu	Glu	Asp	Asp	
		24	25			:	2430				243	5				
acc	acc	gtg	tgc	tgc	tee	atg	tca	tac	tee	tgg	acc	aaa	gct	cta	ata	7699
Thr	Thr	Val	Сув	Сув	Ser	Met	Ser	Tyr	Ser	Trp	Thr	Gly	Ala	Leu	Ile	
	24	40				2445				24	50					
act	ccc	tgt	agc	ccc	gaa	gag	gaa	aag	ttg	cca	atc	aac	cct	ttg	agt	7747
Thr	Pro	Cys	Ser	Pro	Glu	Glu	Glu	Lys	Leu	Pro	Ile	Asn	Pro	Leu	Ser	
2	455				246	0			24	65						
aac	tcg	ctg	ttg	cga	tac	cat	aac	aag	gtg	tac	tgt	aca	aca	tca	aag	7795
Asn	Ser	Leu	Leu	Arg	Tyr	His	Asn	Lys	Val	Tyr	Cys	Thr	Thr	Ser	Lys	
2470	)			247	5			2	480				248	5		
agc	gcc	tca	cag	agg	gct	aaa	aag	gta	act	ttt	gac	agg	acg	caa	gtg	7843
Ser	Ala	Ser	Gln	Arg	Ala	Гуз	Гуз	Val	Thr	Phe	Asp	Arg	Thr	Gln	Val	
			249	0			2	495				2500	)			

ctc	gac	gcc	cat	tat	gac	tca	gtc	tta	aag	gac	atc	aag	cta	gcg	gct	7891
Leu	Asp	Ala	His	Tyr	Asp	Ser	Val	Leu	ГÀв	Asp	Ile	ГÀЗ	Leu	Ala	Ala	
		25	05			2	2510				251	.5				
tcc	aag	gtc	agc	gca	agg	ctc	ctc	acc	ttg	gag	gag	gcg	tgc	cag	ttg	7939
Ser	Lys	Val	ser	Ala	Arg	Leu	Leu	Thr	Leu	Glu	Glu	Ala	Сув	Gln	Leu	
	2 !	520				2525				25	3 0					
act	cca	ccc	cat	tct	gca	aga	tcc	aag	tat	gga	ttc	999	gcc	aag	gag	7987
Thr	Pro	Pro	His	Ser	Ala	Arg	Ser	Lys	Tyr	Gly	Phe	Gly	Ala	Lys	Glu	
2	535				254	0			25	545						
gtc	cgc	agc	ttg	tcc	999	agg	gcc	gtt	aac	cac	atc	aag	tcc	gtg	tgg	8035
Val	Arg	Ser	Leu	Ser	Gly	Arg	Ala	Val	Asn	His	Ile	Lys	ser	Val	Trp	
2550	)			255	55			2	560				256	5		
				~	as a	000	caa	aca	cca	att	ccc	aca	acc	atc	a t. a	8083
aag	gac	ctc	etg	yaa	gac	cca										0003
			Leu						Pro			Thr				0003
			_	Glu			Gln		Pro			Thr 2580	Thr			0003
			Leu	Glu			Gln	Thr	Pro				Thr			0003
Lys	Asp	Leu	Leu	Glu '0	Asp	Pro	Gln 2	Thr 575		Ile	Pro	2586	Thr	Ile	Met	8131
gcc	Asp	Leu	Leu 257	Glu '0 gtg	Asp	Pro tgc	Gln 2 gtg	Thr 575 gac	ccc	Ile gcc	Pro aag	2586 ggg	Thr ) ggt	Ile aag	Met aaa	
gcc	Asp	Leu	Leu 257 gag Glu	Glu '0 gtg	Asp	Pro tgc Cys	Gln 2 gtg	Thr 575 gac	ccc	Ile gcc	Pro aag	2586 ggg	Thr ) ggt	Ile aag	Met aaa	
gcc	Asp	Leu aat Asn	Leu 257 gag Glu	Glu '0 gtg	Asp	Pro tgc Cys	Gln 2 gtg Val	Thr 575 gac	ccc	Ile gcc	Pro aag Lys	2586 ggg	Thr ) ggt	Ile aag	Met aaa	
Lys gcc Ala	Asp aaa Lys	Leu aat Asn 25	Leu 257 gag Glu 85	Glu '0 gtg Val	Asp ttc Phe	Pro tgc Cys	Gln 2 gtg Val	Thr 575 gac Asp	ccc Pro	Ile gcc Ala	Pro aag Lys 259	2586 999 Gly 5	Thr ) ggt Gly	Ile aag Lys	Met aaa	8131
Lys gcc Ala	Asp aaa Lys gct	Leu aat Asn 25	Leu 257 gag Glu 85	Glu 70 gtg Val	Asp ttc Phe	Pro tgc Cys	gtg Val	Thr 575 gac Asp	ccc Pro	gcc Ala	Pro aag Lys 259 gtc	2586 999 Gly 5	Thr ggt Gly gtc	Ile aag Lys	Met aaa Lys gag	8131
Lys gcc Ala	Asp aaa Lys gct Ala	Leu aat Asn 25	Leu 257 gag Glu 85	Glu 70 gtg Val	ttc Phe gtt Val	Pro tgc Cys	gtg Val 2590 cct	Thr 575 gac Asp	ccc Pro	gcc Ala	aag Lys 259 gtc Val	2586 999 Gly 5	Thr ggt Gly gtc	Ile aag Lys	Met aaa Lys gag	8131
Lys gcc Ala	Asp aaa Lys gct Ala	aat Asn 25 cgc	Leu 257 gag Glu 85	Glu 70 gtg Val	ttc Phe gtt Val	tgc Cys tac Tyr	gtg Val 2590 cct	Thr 575 gac Asp	ccc Pro	gcc Ala ggc Gly	aag Lys 259 gtc Val	2586 999 Gly 5	Thr ggt Gly gtc	Ile aag Lys	Met aaa Lys gag	8131
Lys gcc Ala cca	Asp aaa Lys gct Ala	aat Asn 25 cgc Arg	Leu 257 gag Glu 85	Glu 70 gtg Val atc	ttc Phe gtt Val	tgc Cys tac Tyr 2605	gtg Val 2590 cct Pro	Thr 575 gac Asp gac Asp	ccc Pro ctc Leu	gcc Ala ggc Gly 26	aag Lys 259 gtc Val	ggg Gly 5 cgg Arg	ggt Gly gtc Val	aag Lys tgc Cys	Met aaa Lys gag Glu	8131
Lys gcc Ala cca Pro	Asp aaa Lys gct Ala 20	aat Asn 25 cgc Arg	Leu 257 gag Glu 85 ctc Leu	Glu  Gly  Gtg  Val  atc  Ile	ttc Phe gtt Val	tgc Cys tac Tyr 2605	gtg Val 2590 cct Pro	Thr 575 gac Asp gac Asp	ccc Pro ctc Leu	gcc Ala ggc Gly 26:	aag Lys 259 gtc Val	ggg Gly 5 cgg Arg	ggt Gly gtc Val	aag Lys tgc Cys	Met  aaa Lys  gag Glu  atg	8131 8179

gga	gct	tee	tat	ggc	ttc	cag	tac	tee	cct	gcc	caa	cgg	gtg	gag	tat	8275
Gly	Ala	Ser	Tyr	Gly	Phe	Gln	Tyr	Ser	Pro	Ala	Gln	Arg	Val	Glu	Tyr	
2630	)			263	5			2	640				2649	5		
ctc	ttg	aaa	gca	tgg	gcg	gaa	aag	aag	gac	ccc	atg	ggt	ttt	tcg	tat	8323
Leu	Leu	Lys	Ala	Trp	Ala	Glu	Гуз	Lys	Asp	Pro	Met	Gly	Phe	Ser	Tyr	
			265	0			2	655				2660	)			
gat	acc	cga	tgc	ttc	gac	tca	acc	gtc	act	gag	aga	gac	atc	agg	acc	8371
qaA	Thr	Arg	Cys	Phe	Asp	Ser	Thr	Va1	Thr	Glu	Arg	Asp	Ile	Arg	Thr	
		26	65			2	2670				267	5				
gag	gag	tcc	ata	tac	cag	gcc	tgc	tec	ctg	ccc	gag	gag	gcc	cgc	act	8419
Glu	Glu	Ser	Ile	Tyr	Gln	Ala	Сув	Ser	Leu	Pro	Glu	Glu	Ala	Arg	Thr	
	26	80				2685				26	90					
gcc	ata	cac	tcg	ctg	act	gag	aga	ctt	tac	gta	gga	999	ccc	atg	ttc	8467
Ala	Ile	His	Ser	Leu	Thr	Glu	Arg	Leu	Тух	Val	Gly	Gly	Pro	Met	Phe	
2	695				270	0			27	05						
aac	agc	aag	ggt	caa	acc	tgc	ggt	tac	aga	cgt	tgc	ege	gcc	agc	999	8515
Asn	Ser	Lys	Gly	Gln	Thr	Cys	Gly	Tyr	Arg	Arg	Cys	Arg	Ala	Ser	Gly	
2710	)			271	.5			2	720				272	5		
gtg	cta	acc	act	agc	atg	ggt	aac	acc	atc	aca	tgc	tat	gtg	aaa	gcc	8563
Val	Leu	Thr	Thr	Ser	Met	Gly	Asn	Thr	Ile	Thr	Cys	Tyr	Val	Lys	Ala	
			273	0			2	735				2740	)			
cta	gcg	gcc	tgc	aag	gct	gcg	999	ata	gtt	gcg	ccc	aca	atg	ctg	gta	8611
Leu	Ala	Ala	Сув	Lys	Ala	Ala	Gly	Ίle	Val	Ala	Pro	Thr	Met	Leu	Val	
		27	45			2	750				275	5				

tgc	ggc	gat	gac	cta	gta	gtc	atc	tca	gaa	agc	cag	aaa	act	gag	gag	8659
Сув	Gly	Asp	Asp	Leu	Val	Val	Ile	Ser	Glu	Ser	Gln	Gly	Thr	Glu	Glu	
	2	760				2765				27	70					
gac	gag	cgg	aac	ctg	aga	gcc	ttc	acg	gag	gcc	atg	acc	agg	tac	tat	8707
Asp	Glu	Arg	Asn	Leu	Arg	Ala	Phe	Thr	Glu	Ala	Met	Thr	Arg	Tyr	Ser	
2	775				278	0			2.	85						
gac	act	cct	ggt	gat	aaa	ccc	aga	aag	gaa	tat	gac	ctg	gag	cta	ata	8755
Ala	Pro	Pro	Gly	Asp	Pro	Pro	Arg	Pro	Glu	Tyr	Asp	Leu	Glu	Leu	Ile	
279	0			279	5			2	800				280	5		
aça	tec	tgt	tcc	tca	aat	gtg	tct	gtg	gcg	ttg	ggc	ccg	cgg	ggc	cgc	8803
Thr	Ser	Cys	Ser	Ser	Asn	Val	Ser	Val	Ala	Leu	Gly	Pro	Arg	Gly	Arg	
			281	. 0			2	815				282	0			
cgc	aga	tac	tac	ctg	acc	aga	gac	cca	acc	act	cca	ctc	gcc	cgg	gct	8851
Arg	Arg	Tyr	Туг	Leu	Thr	Arg	Asp	Pro	Thr	Thr	Pro	Leu	Ala	Arg	Ala	
		28	25			2	2830				283	5				
gec	tgg	gaa	aca	gtt	aga	cac	tcc	act	atc	aat	tca	tgg	ctg	gga	aac	8899
Ala	Trp	Glu	Thr	Val	Arg	His	Ser	Pro	Ile	Asn	Ser	Trp	Leu	Gly	Asn	
	28	340				2845				28	50					
atc	atc	cag	tat	gct	cca	acc	ata	tgg	gtt	cgc	atg	gtc	cta	atg	aca	8947
Ile	Ile	Gln	Tyr	Ala	Pro	Thr	Ile	Trp	Val	Arg	Met	Val	Leu	Met	Thr	
2	855				2860	0			28	65						
cac	ttc	ttc	tcc	att	ctc	atg	gtc	caa	gac	acc	ctg	gac	cag	aac	ctc	8995
His	Phe	Phe	Ser	Ile	Leu	Met	Val	Gln	Asp	Thr	Leu	Asp	Gln	Asn	Leu	
2870	)			287	5			2	880				2885	<b>;</b>		

aac	ttt	gag	atg	tat	gga	tca	gta	tac	tcc	gtg	aat	cct	ttg	gac	ctt	9043
Asn	Phe	Glu	Met	Tyr	Gly	Ser	Va1	Tyr	Ser	Val	Asn	Pro	Leu	Asp	Leu	
			289	90			2	895				290	)			
cca	gcc	ata	att	gag	agg	tta	caç	ggg	ctt	gac	gcc	ttt	tct	atg	cac	9091
Pro	Ala	Ile	Ile	Glu	Arg	Leu	His	Gly	Leu	Asp	Ala	Phe	Ser	Met	His	
		29	05			2	2910				291	5				
aca	tac	tct	cac	cac	gaa	ctg	acg	cgg	gtg	g¢t	tça	gcc	ctc	aga	aaa	9139
Thr	Tyr	Ser	His	His	Glu	Leu	Thr	Arg	Val	Ala	Ser	Ala	Leu	Arg	Lys	
	25	920				2925				29	3 0					
ctt	aaa	gcg	cca	ccc	ctc	agg	gtg	tgg	aag	agt	cgg	gct	cgc	gca	gtc	9187
Leu	Gly	Ala	Pro	Pro	Leu	Arg	Val	Trp	Lys	Ser	Arg	Ala	Arg	Ala	Val	
2	935				294	0			29	145						
agg	gcg	tec	ctc	atc	tcc	cgt	gga	ggg	aaa	gcg	gcc	gtt	tgc	ggc	cga	9235
Arg	Ala	Ser	Leu	Ile	ser	Arg	Gly	Gly	Lys	Ala	Ala	Val	Cys	Gly	Arg	
2950	)			295	5			2	960				296	5		
tat	ctc	ttc	aat	tgg	gcg	gtg	aag	acc	aag	ctc	aaa	ctc	act	cca	ttg	9283
Tyr	Leu	Phe	Asn	Trp	Ala	Val	Lys	Thr	Lys	Leu	Lуs	Leu	Thr	Pro	Leu	
			297	0			2	975				2980	)			
ccg	gag	gcg	cgc	cta	ctg	gac	tta	tcc	agt	tgg	ttc	acc	gtc	ggc	gcc	9331
Pro	Glu	Ala	Arg	Leu	Leu	Asp	Leu	Ser	Ser	Trp	Phe	Thr	Val	Gly	Ala	
		29	85			2	990				299	5				
ggc	ggg	ggc	gac	att	ttt	cac	agc	gtg	teg	aga	gcc	cga	ccc	cgc	tca	9379
Gly	Gly	Gly	Asp	Ile	Phe	His	Ser	Val	Ser	Arg	Ala	Arg	Pro	Arg	Ser	
	3 0	000				3005				3 0 3	10					

tta ctc ttc ggc cta ctc cta ctt ttc gta ggg gta ggc ctc ttc cta 9427 Leu Leu Phe Gly Leu Leu Leu Phe Val Gly Val Gly Leu Phe Leu 3015 3020 3025

ctc ccc gct cgg tag agcggcacac actaggtaca ctccatagct aactgttcct 9482 Leu Pro Ala Arg

3030

tcacggctag ctgtgaaagg tccgtgagcc gcatgactgc agagagtgcc gtaactggtc 9662

tctctgcaga tcatgt

<210> 4

<211> 3033

<212> PRT

<213> Hepatitis C virus

<400> 4

Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn

1 5 10 15

Arg Arg Pro Glu Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly

0 25 30

Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Thr

5 40 4

Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro

50 55 60

Ile	Pro	Lys	Asp	Arg	Arg	Ser	Thr	Gly	Lys	Ala	ттр	Gly	Гув	Pro	Gly
65				7	0				75				80		
Arg	Pro	Trp	Pro	Leu	Tyr	Gly	Asn	Glu	Gly	Leu	Gly	Trp	Ala	Gly	Trp
			8	5				90				95			
Leu	Leu	Ser	Pro	Arg	Gly	Ser	Arg	Pro	Ser	Trp	Gly	Pro	Thr	Asp	Pro
		1	00				105				11	0			
Arg	His	Arg	Ser	Arg	Asn	Va1	Gly	Lys	Val	Ile	Asp	Thr	Leu	Thr	Сув
	1	.15				120				12	5				
Gly	Phe	Ala	Asp	Leu	Met	Gly	Tyr	Ile	Pro	Val	Val	Gly	Ala	Pro	Leu
	130				135	5			1	4 0					
Ser	Gly	Ala	Ala	Arg	Ala	Val	Ala	His	Gly	۷al	Arg	Val	Leu	Glu	Asp
145				15	0			:	155				160		
Gly	Val	Asn	Tyr	Ala	Thr	Gly	Asn	Leu	Pro	Gly	Phe	Pro	Phe	Ser	Ile
			16	5				170				175			
Phe	Leu	Leu	Ala	Leu	Leu	Ser	Cys	Ile	Thr	Val	Pro	Val	Ser	Ala	Ala
		1	80				185				19	0			
Gln	Val	Lys	Asn	Thr	Ser	Ser	Ser	Tyr	Met	Val	Thr	Asn	Asp	Cys	Ser
	1	.95				200				20	5				
Asn	Asp	Ser	Ile	Thr	тхр	Gln	Leu	Glu	Ala	Ala	Val	Leu	His	Val	Pro
;	210				215	•			2	20					
Gly	Cys	Val	Pro	Сув	Glu	Arg	Val	Gly	neA	Thr	Ser	Arg	Сув	Trp	Val
225				23	0			2	35				240		
Pro	Val	Ser	Pro	Asn	Met	Ala	Val	Arg	Gln	Pro	Gly	Ala	Leu	Thr	Gln
			24	5			2	250				255			
Gly	Leu	Arg	Thr	His	Ile	Asp	Met	Val	Val	Met	ser	Ala	Thr	Phe	Сув
		20	60				265				270	)			
Ser	Ala	Leu	Tyr	Val	Gly	Asp	Leu	Cys	Gly	Gly	Va1	Met	Leu	Ala	Ala
	2	75				280				28	5				
Gln	Val	Phe	Ile	Va1	ser	Pro	Gln	Tyr	His	Trp	Phe	Val	Gln	Glu	Сув
:	290				295				3	00					
Asn	Cys	Ser	Ile	Tyr	Pro	Gly	Thr	Ile	Thr	Gly	нія	Arg	Met	Ala	Trp
305				31:	0			7	115				320		

Asp	Met	Met	Met	Asn	Trp	Ser	Pro	Thr	Ala	Thr	Met	Ile	Leu	Ala	Tyr
			3 2	5			;	330				335			
Val	Met	Arg	Val	Pro	Glu	Val	Ile	Ile	Asp	Ile	Va1	Ser	Gly	Ala	His
		3	40				345				35	0			
Trp	Gly	Val	Met	Phe	Gly	Leu	Ala	Tyr	Phe	Ser	Met	Gln	Gly	Ala	Trp
	3	55				360				36	55				
Ala	Lys	Val	Ile	Val	Ile	Leu	Leu	Leu	Ala	Ala	Gly	Val	Asp	Ala	Gly
	370				375	5			3	80					
Thr	Thr	Thr	Val	Gly	Gly	Ala	Val	Ala	Arg	Ser	Thr	Asn	Val	Ile	Ala
385				3 9	0			3	95				400		
Gly	Val	Phe	Ser	His	Gly	Pro	Gln	Gln	Asn	Ile	Gln	Leu	Ile	Asn	Thr
			40	5				410				415			
Asn	Gly	Ser	Trp	His	Ile	Asn	Arg	Thr	Ala	Leu	Asn	Cys	Asn	Asp	Ser
		4	20				425				43	0			
Leu	Asn	Thr	Gly	Phe	Leu	Ala	Ala	Leu	Phe	Tyr	Thr	Asn	Arg	Phe	Asn
	4	35				440				44	. 5				
Ser	Ser	Gly	Сув	Pro	Gly	Arg	Leu	Ser	Ala	Сув	Arg	Asn	Ile	Glu	Ala
	450				455	5			4	60					
Phe	Arg	Ile	Gly	Trp	Gly	Thr	Leu	Gln	Tyr	Glu	Asp	Asn	Val	Thr	Asn
465				47	0			4	175				480		
Pro	Glu	Asp	Met	Arg	Pro	Tyr	Сув	Trp	His	Tyr	Pro	Pro	Lys	Pro	Сув
			48	5			•	190				495			
Gly	Val	Val	Pro	Ala	Arg	Ser	Val	Cys	Gly	Pro	Va l	Tyr	Cys	Phe	Thr
		5	00				505				51	0			
Pro	Ser	Pro	Val	Val	Val	Gly	Thr	Thr	Asp	Arg	Arg	Gly	Val	Pro	Thr
	5	15				520				52	5				
Тук	Thr	Trp	Gly	Glu	Asn	Glu	Thr	Asp	Val	Phe	Leu	Leu	Asn	Ser	Thr
	530				535	i			5	4 0					
Arg	Pro	Pro	Gln	Gly	Ser	Trp	Phe	Gly	Сув	Thr	Trp	Met	Asn	Ser	Thr
545				55	0			9	555				560		
Gly	Phe	Thr	Lys	Thr	Cys	Gly	Ala	Pro	Pro	Сув	Arg	Thr	Arg	Ala	Asp
			E 6	e				70				575			

Phe	Asn	Ala	ser	Thr	Asp	Leu	Leu	Cys	Pro	Thr	Asp	Cys	Phe	Arg	Lys
		5	80				585				59	0			
His	Pro	Asp	Ala	Thr	Tyr	Ile	Lys	Cys	Gly	Ser	Gly	Pro	Trp	Leu	Thr
	5	95				600				60	5				
Pro	Lys	Cys	Leu	Val	His	Tyr	Pro	Tyr	Arg	Leu	Trp	His	Tyr	Pro	Сув
	610				615	i			6	20					
Thr	Val	Asn	Phe	Thr	Ile	Phe	Lys	Ile	Arg	Met	Туг	Val	Gly	Gly	Val
625				63	0			6	535				640		
Glu	His	Arg	Leu	Thr	Ala	Ala	Cys	Asn	Phe	Thr	Arg	Gly	Asp	Arg	Сув
			64	5			i	650				655			
Asp	Leu	Glu	Asp	Arg	Asp	Arg	Ser	Gln	Leu	Ser	Pro	Leu	Leu	His	Ser
		6-	60				665				67	0			
Thr	Thr	Glu	Trp	Ala	Ile	Leu	Pro	Сув	Thr	Tyr	Ser	Asp	Leu	Pro	Ala
	6	75				680				68	3 5				
Leu	ser	Thr	Gly	Leu	Leu	His	Leu	His	Gln	Asn	Ile	Val	Asp	Val	Gln
	690				695	5			7	00					
Tyr	Met	Tyr	Gly	Leu	Ser	Pro	Ala	Ile	Thr	Lys	Tyr	Val	Val	Arg	Trp
705				71	0			-	715				720		
Glu	Trp	Val	Val	Leu	Leu	Phe	Leu	Leu	Leu	Ala	Asp	Ala	Arg	Val	Сув
			72	5			,	730				735			
Ala	Сув	Leu	Trp	Met	Leu	Ile	Leu	Leu	Gly	Gln	Ala	Glu	Ala	Ala	Leu
		7	40				745				75	0			
Glu	Гуз	Leu	Val	Val	Leu	His	Ala	Ala	Ser	Ala	Ala	Asn	Cys	His	Gly
	7	755				760				76	5 5				
Leu	Leu	Tyr	Phe	Ala	Ile	Phe	Phe	Val	Ala	Ala	Trp	His	Ile	Arg	Gly
	770				775	5			7	80					
Arg	Val	Val	Pro	Leu	Thr	Thr	Tyr	Сув	Leu	Thr	Gly	Leu	Trp	Pro	Phe
785				79	0			-	795				800		
Сув	Leu	Leu	Leu	Met	Ala	Leu	Pro	Arg	Gln	Ala	Tyr	Ala	Tyr	Asp	Ala
			80	5			:	810				815			
Pro	Val	Нis	Gly	Gln	Ile	Gly	Val	Gly	Leu	Ьeu	Ile	Leu	Ile	Thr	Leu
		8	20				825				83	0			

Phe	Thr	Leu	Thr	Pro	Gly	туг	Lys	Thr	Leu	Leu	Gly	Gln	Cys	Leu	Trp
	8	3 5				840				84	5				
Trp	Leu	Сув	Tyr	Leu	Leu	Thr	Leu	Gly	Glu	Ala	Met	Ile	Gln	Glu	Trp
	850				855	,			8	60					
Val	Pro	Pro	Met	Gln	Va1	Arg	Gly	Gly	Arg	Asp	Gly	Ile	Ala	Trp	Ala
865				87	0			8	375				880		
Val	Thr	Ile	Phe	Сув	Pro	Gly	Va1	Val	Phe	Asp	Ile	Thr	Lys	Trp	Leu
			88	5			\$	890				895			
Leu	Ala	Leu	Leu	Gly	Pro	Ala	Tyr	Leu	Leu	Arg	Ala	Ala	Leu	Thr	His
		9	00				905				91	0			
Val	Pro	Tyr	Phe	Val	Arg	Ala	His	Ala	Leu	Ile	Arg	۷al	Сув	Ala	Leu
	9	15				920				92	25				
Val	Ъуѕ	Gln	Leu	Ala	Gly	Gly	Arg	Tyr	Val	Gln	Val	Ala	Leu	Leu	Ala
	930				935					40					
Leu	Gly	Arg	Trp	Thr	Gly	Thr	Tyr	Ile	туг	Asp	His	Leu	Thr	Pro	Met
945				95					955				960		
Ser	Asp	Trp	Ala	Ala	Ser	Gly	Leu	Arg	Asp	Leu	Ala	Val	Ala	Val	Glu
			96					970				975			
Pro	Ile	Ile	Phe	Ser	Pro	Met	Glu	Lys	Lys	Val	ıle	Val	Trp	Gly	Ala
			80				985				99				
Glu	Thr	Ala	Ala	Cys	Gly	Asp	Ile	Leu	His	Gly	Leu	Pro	Val	Ser	Ala
		995				1000				10					
Arg	Leu	Gly	Gln	Glu	Ile	Leu	Leu	Gly			Asp	Gly	Tyr	Thr	Ser
	010				101					020					
Lys	Gly	Trp	Lys	Leu	Leu	Ala	Pro			Ala	Tyr	Ala			Thr
102				103					035		_	_	104		_
Arg	Gly	Leu		Gly	Ala	Ile			Ser	Met	Thr			Asp	Arg
			104			_		050		_	_,	105		<b>a</b> 1 .	
Thr	Glu			Gly	Glu			Ile	Leu	Ser			ser	GID	ser
	_		60				1065	** - 7	•	m	10		m	11 i =	01
Phe			Thr	Thr	lle			val	ьeu			val	тут	HIS	стА
	1	075				1080	,			ŦÜ	85				

Ala	Gly	Asn	Lys	Thr	Leu	Ala	Gly	Leu	Arg	Gly	Pro	Val	Thr	Gln	Met
	090				1095				11						
TVY	Ser	Ser	Ala	Glu	Gly	Asp	Leu	Val	Gly	Trp	Pro	Ser	Pro	Pro	Gly
1109				111					115				1120		
Thr	LVS	Ser	Leu	Glu	Pro	Cys	Lys	Сув	Gly	Ala	Val	Asp	Leu	Tyr	Leu
1111	цуо		11:					130				1135			
7703	mh r	ልተጠ		Ala	qeA	٧al	Ile	Pro	Ala	Arg	Arg	Arg	Gly	qaA	ГÀв
vai	1117		.40		_		L145				115				
3	a su			Leu	Ser	Pro	Arg	Pro	Ile	Ser	Thr	Leu	Lуя	Gly	ser
Arg			неи			1160					65				
_		155	. Bec	val	ĩ.en			Arq	Gly	His	Val	Val	Gly	Leu	Phe
		СТУ	PIC	, va.	117			_		180					
	L170 -			0			- GTv	Val	Ala	Lys	Ser	Ile	Asp	Phe	Ile
Arg	Ala	Ala	ı val			n. n.	U11		1195	•			120		
118	5				.90		T e W			. Ser	Pro	Thr	Phe	Ser	Asp
Pro	Val	L Glu			ı ASI	) Vai		1210				121			
				205		3			. መሉን	- ጥህን	r Gli			туз	. Leu
Asr	. Se	r Th:	r Pr	o Pro	o Ala	a Val			1 1111	. 19.		30	1	-	: Leu
			220				1225			<b>.</b>			o Val	A1;	a Tvr
His	a Ala	a Pr	o Th	r Gly	y Se:	r Gly	у Бу:	s Se	r yni			L FIV	<i>,</i> vu.		a Tyr
		1235				124					245	- <b>G</b> a-	∝ Vn`	ו או.	a Ala
Ala	a Al	a Gl	n Gl	у Ту	r Ly	s Val	l Le	u Va			n Pr	o se.	r Va.	LAL	a Ala
	1250					55				1260			_1		Dwo
Th	r Le	u Gl	y Ph	ie Gl	y Al	а Ту	r Le	u Se	т Ьу	s Al	a Hi	s G1			n Pro
12	65				270				1275				12		1 -
As	n Il	e Ar	g Tł	ır Gl	y Va	1 Ar	g Th	r Va	1 Me	t Th	r Gl	y Gl	u Al	a Il	e Thr
				285				129					95		_
Ту	r Se	r Tì	ır Ty	r Gl	ry ry	s Ph	e Le	u Al	a As	p G1	y G1	у Су	g Al	a Se	r Gly
			1300				130					310			
Al	аТу	r A	sp I	le II	le I1	Le Cy	rs As	sp G	lu Cy	s Hi	is Al	La Va	ıl As	p Al	la Thr
		131	5			13	20			;	1325				
Se	er II	le L	eu G	ly I	le G	ly Th	ır Va	al Le	eu As	sp G	ln A	La G	lu Tł	ır A.	la Gly
	122					335				1340					

Val	Arg	Leu	Thr	Val	Leu	Ala	Thr	Ala	Thr	Pro	Pro	Gly	Ser	Val	Thr
1345	5			135	0			1	355				1360	)	
Thr	Pro	His	Pro	Asp	Ile	Glu	Glu	Val	Gly	Leu	Gly	Arg	Glu	Gly	Glu
			136	5			1	370				1375	5		
Ile	Pro	Phe	Tyr	Gly	Arg	Ala	Ile	Pro	Leu	Ser	Сув	Ile	Гуз	Gly	Gly
		13	80			1	385				139	0			
Arg	His	Leu	Ile	Phe	Cys	His	Ser	Lys	Lys	Lys	Cys	Asp	Glu	Leu	Ala
	13	95				1400				14	05				
Ala	Ala	Leu	Arg	Gly	Met	Gly	Leu	Asn	Ala	Val	Ala	Tyr	Tyr	Arg	Gly
1	410				141	5			14	20					
Leu	Asp	Val	ser	Ile	Ile	Pro	Ala	Gln	сіу	Asp	Val	Val	Va1	Val	Ala
1425	5			143	0			1	435				1440	)	
Thr	Asp	Ala	Leu	Met	Thr	Gly	Tyr	Thr	Gly	Asp	Phe	Asp	Ser	Val	Ile
			144	5			i	450				1455	5		
Asp	Cys	Asn	Val	Ala	Val.	Thr	Gln	Ala	Va1	Asp	Phe	Ser	Leu	Asp	Pro
		14	60			1	L465				147	0			
Thr	Phe	Thr	Ile	Thr	Thr	Gln	Thr	Val	Pro	Gln	Asp	Ala	Va1	Ser	Arg
	14	175				1480				14	85				
Ser	Gln	Arg	Arg	Gly	Arg	Thr	Gly	Arg	Gly	Arg	Gln	Gly	Thr	Tyr	Arg
1	490				149	5			19	00					
Tyr	Val	Ser	Thr	Gly	Glu	Arg	Ala	Ser	Gly	Met	Phe	Asp	Ser	Val	Val
150	5			151	. 0			1	515				152	0	
Leu	Сув	Glu	Cys	Tyr	Asp	Ala	Gly	Ala	Ala	Trp	Tyr	Asp	Leu	Thr	Pro
			152	:5			1	530				153	5		
Ala	Glu	Thr	Thr	Val	Arg	Leu	Arg	Ala	Tyr	Phe	Asn	Thr	Pro	Gly	Leu
		15	40			1	1545				155	0			
Pro	Val	Сув	Gln	Asp	His	Leu	Glu	Phe	Trp	Glu	Ala	Val	Phe	Thr	Gly
	15	555				1560				15	65				
Leu	Thr	His	Ile	Asp	Ala	His	Phe	Leu	Ser	Gln	Thr	Lys	Gln	Ala	Gly
1	570				157	5			15	80					
Glu	Asn	Phe	Ala	Tyr	Leu	Val	Ala	Туг	Gln	Ala	Thr	Val	Сув	Ala	Arg
1589	5			159	0 0			1	595				160	0	

Ala	Гуs	Ala	Pro	Pro	Pro	Ser	Trp	Asp	Ala	Met	Trp	Lys	Сув	Leu	Ala
			160	5			1	610				1615	5		
Arg	Leu	ГÀа	Pro	Thr	Leu	Ala	Gly	Pro	Thr	Pro	Leu	Leu	Тух	Arg	Leu
		16	20			1	625				163	0			
Gly	Pro	Ile	Thr	Asn	Glu	Val	Thr	Leu	Thr	His	Pro	Gly	Thr	Lys	Tyr
	16	35				1640				16	45				
Ile	Ala	Thr	Cys	Met	Gln	Ala	Asp	Leu	Glu	Val	Met	Thr	Ser	Thr	Trp
1	650				165	5			1.6	60					
Val	Leu	Ala	Gly	Gly	Va 1	Leu	Ala	Ala	Val	Ala	Ala	Tyr	Сув	Leu	Ala
1665	5			167	0			1	675				1680	)	
Thr	Gly	Сув	Val	Ser	Ile	Ile	Gly	Arg	Leu	His	Val	Asn	Gln	Arg	Val
			168	15			1	690				1699	5		
Val	Val	Ala	Pro	qeA	Lys	Glu	Val	Leu	Tyr	Glu	Ala	Phe	Asp	G1u	Met
		17	00			-	1705				171	0			
Glu	Glu	Сув	Ala	Ser	Arg	Ala	Ala	Leu	Ile	Glu	Glu	Gly	Gln	Arg	Ile
	13	715				1720				17	25				
Ala	Glu	Met	Leu	Lys	Ser	Lys	Ile	Gln	Gly	Leu	Leu	Gln	Gln	Ala	Ser
1	730				173	5			17	740					
Lys	Gln	Ala	Gln	Asp	Ile	Gln	Pro	Ala	Met	Gln	Ala	Ser	Trp	Pro	Lys
1749	i			175	0			1	755				176	D	
Val	Glu	Gln	Phe	Trp	Ala	Arg	His	Met	ттр	Asn	Phe	Ile	Ser	Gly	Ile
			176	55			1	770				1779	5		
Gln	Tyr	Leu	Ala	Gly	Leu	Ser	Thr	Leu	Pro	Gly	Asn	Pro	Ala	Val	Ala
		17	80			-	1785				179	0			
Ser	Met	Met	Ala	Phe	Ser	Ala	Ala	Leu	Thr	ser	Pro	Leu	Ser	Thr	Ser
	1′	795				1800				18	05				
Thr	Thr	Ile	Leu	Leu	Asn	Ile	Met	Gly	Gly	Trp	Leu	Ala	Ser	Gln	Ile
1	810				181	5			1.8	320					
Ala	Pro	Pro	Ala	Gly	Ala	Thr	Gly	Phe	Val	Val	Ser	Gly	Leu	Val	Gly
1829	Pro Pro Ala Gly Ala 5 5 1830							1.	835				184	0	
Ala	Ala	Val	Gly	Ser	Ile	Gly	Leu	Gly	Lys	Val	Lеи	Va1	Asp	Ile	Leu
			184	15			1	850				185	5		

Ala	Gly	Tyr	Gly	Ala	Gly	Ile	Ser	Gly	Ala	Leu	Val	Ala	Phe	Lys	Ile
		18	60			3	865				187	0			
Met	ser	Gly	Glu	Lys	Pro	Ser	Met	Glu	Asp	Val	Ile	Asn	Leu	Leu	Pro
	18	375				1880				188	85				
Gly	Ile	Leu	Ser	Pro	Gly	Ala	Leu	Val	Val	Gly	Val	Ίle	Cys	Ala	Ala
1	890				189	5			1.9	00					
Ile	Leu	Arg	Arg	His	Val	Gly	Pro	Gly	Glu	Gly	Ala	Val	Gln	Trp	Met
1909	õ			191	. 0			1	915				1920	)	
Asn	Arg	Leu	Ile	Ala	Phe	Ala	Ser	Arg	Gly	Asn	His	Va1	Ala	Pro	Thr
			192	:5			1	930				1935	5		
His	Tyr	Val	Thr	Glu	ser	Asp	Ala	Ser	Gln	Arg	Va1	Thr	Gln	Leu	Leu
		19	40				1945				195	0			
Gly	Ser	Leu	Thr	Ile	Thr	Ser	Leu	Leu	Arg	Arg	Leu	His	Asn	Trp	Ile
	19	955				1960				19	65				
Thr	Glu	Asp	Cys	Pro	Ile	Pro	Сув	Ser	Gly	Ser	Trp	Leu	Arg	Asp	Val
1	970				197	5			19	08					
Trp	Asp	Trp	Val	Сув	Thr	Ile	Leu	Thr	Asp	Phe	ГЛа	Asn	Trp	Leu	Thr
198	5			199	90			1	995				200	D	
Ser	Lys	Leu	Phe	Pro	Lys	Leu	Pro	Gly	Leu	Pro	Phe	Ile	Ser	Сув	Gln
			200	5			2	010				201	5		
Lys	Gly	Туг	Lys	Gly	Val	Trp	Ala	Gly	Thr	Gly	Ile	Met	Thr	Thr	Arg
		20	20			:	2025				203	0			
Cys	Pro	Cys	Gly	Ala	Asn	Ile	Ser	Gly	Asn	Val	Arg	Lец	Gly	Ser	Met
	.20	335				2040	ı			20	4 5				
Arg	Ile	Thr	Gly	Pro	Lys	Thr	Сув	Met	Asn	Thr	Trp	Gln	Gly	Thr	Phe
2	050				205	5			20	060					
Pro	Ile	Asn	Сув	Туг	Thr	Glu	Gly	Gln	Cys	Ala	Pro	Lys	Pro	Pro	Thr
206	ő			207	70			2	075				208	0	
Asn	Tyr	ГАЗ	Thr	Ala	Ile	Trp	Arg	Val	Ala	Ala	Ser	Glu	Tyr	Ala	Glu
			208	35			2	090				209	5		
17a l		0.7	** 1	<b>41</b>	<b>~</b>			ш	11-7	m b se	<b>03.</b>	т	mb		7 0 0
VUL	Thr	Gin	нія	GIY	ser	TYT	ser	туг	Agt	1117	GIĀ	Leu	THE	Thr	Asp

Asn	Leu	Lys	Ile	Pro	Сув	Gln	Leu	Pro	ser	Pro	Glu	Phe	Phe	Ser	Trp
	21	15				2120				212	:5				
Val	Asp	Gly	Val	Gln	Ile	His	Arg	Phe	Ala	Pro	Thr	Pro	Lys	Pro	Phe
2	130				213	5			21	40					
Phe	Arg	Asp	Glu	Val	ser	Phe	Cys	Val	Gly	Leu	Asn	Ser	Tyr	Ala	Val
2145	5			215	0			2	155				2160	)	
Gly	ser	Gln	Leu	Pro	Cys	Glu	Pro	Glu	Pro	Asp	Ala	Asp	Val	Leu	Arg
			216	55			2	170				2175	5		
ser	Met	Leu	Thr	Asp	Pro	Pro	His	Ile	Thr	Ala	Glu	Thr	Ala	Ala	Arg
		21	80			:	2185				219	0			
Arg	Leu	Ala	Arg	Gly	Ser	Pro	Pro	Ser	Glu	Ala	Ser	Ser	Ser	Val	Ser
	2	195				2200	)			22	0 5				
Gln	Leu	Ser	Ala	Pro	Ser	Leu	Arg	Ala	Thr	Cys	Thr	Thr	His	Ser	Asn
2	210				221	5			23	220					
Thr	Tyr	Asp	Val	Asp	Met	Val	Asp	Ala	Asn	Leu	Leu	Met	Glu	Gly	Gly
222	5			22	3 0			2	235				224	0	
Va1	Ala	Gln	Thr	Glu	Pro	Glu	Ser	Arg	Val	Pro	Val	Leu	Asp	Phe	Leu
			22	45			2	250				225	5		
Glu	Pro	Met	Ala	Glu	Glu	Glu	. ser	Asp	Leu	Glu	Pro	ser	Ile	Pro	Ser
		22	260				2265				22	70			
Glu	Суз	Met	Leu	Pro	Arg	Ser	Gly	Phe	Pro	Arg	Ala	Leu	Pro	Ala	Trp
	2	275				228	0			22	85				
Ala	Arg	Pro	Asp	туг	Asn	Pro	Pro	Leu	Val	Glu	Ser	Trp	Arg	Arg	Pro
:	2290				229	95			2	300					
Asp	Туг	Gln	Pro	Pro	Thr	. Val	Ala	ı Gly	Сув	Ala	Leu	Pro	Pro	Pro	Lys
230	5			23	10			:	2315				232	0	
Lys	Ala	ı Pro	Thr	Pro	Pro	Pro	Arg	g Arg	Arg	g Arg	Thr	Val	Gly	Lev	Ser
			23	25				2330				233	5		
Gli	ı Sei	c Thi	: I1e	e Sei	Gli	ı Ala	a Lev	ı Glı	ı Glı	ı Let	Ala	ıle	Lys	Thi	Phe
		2	340				234	5			23	50			
Gly	/ Gl:	ı Pro	o Pro	Sei	s Sei	c Gly	y Asj	alA c	ı Gly	, Sei	Sez	Thr	: Gly	/ Ala	ı Gly
		2355				236					365				

Ala	Ala	Glu	Ser	Gly	Gly	Pro	Thr	Ser	Pro	Gly	Glu	Pro	Ala	Pro	Ser
2	370				237	5			23	80					
3lu	Thr	Gly	Ser	Ala	Ser	Ser	Met	Pro	Pro	Leu	Glu	Gly	Glu	Pro	Gly
2385	5			239	0			2	395				2400	)	
Asp	Pro	Asp	Leu	Glu	Ser	Asp	Gln	Val	Glu	Leu	Gln	Pro	Pro	Pro	Gln
			240	15			2	410				2419	5		
Gly	Gly	Gly	Val	Ala	Pro	Gly	Ser	Gly	ser	Gly	Ser	Trp	Ser	Thr	Сув
		24	20			2	425				243	0			
Ser	Glu	Glu	Asp	Asp	Thr	Thr	Val	Сув	Сув	Ser	Met	Ser	Tyr	Ser	Trp
	24	35				2440				24	4 5				
Thr	Gly	Ala	Leu	Ile	Thr	Pro	Сув	Ser	Pro	Glu	Glu	Glu	Lys	Leu	Pro
2	450				245	5			24	160					
Ile	Asn	Pro	Leu	Ser	Asn	ser	Leu	Leu	Arg	Tyr	His	Asn	Lys	Val	Tyr
246	5			247	70			2	475				248	)	
Cys	Thr	Thr	ser	Lys	Ser	Ala	Ser	Gln	Arg	Ala	Lув	ГÀа	Val	Thr	Phe
			248	3 5			2	490				249	5		
Asp	Arg	Thr	Gln	Val	Leu	Asp	Ala	His	Tyr	Asp	Ser	Val	Leu	Ъуs	Asp
		25	00			:	2505				251	.0			
Ile	Lys	Leu	Ala	Ala	Ser	Lys	Val	Ser	Ala	Arg	Leu	Leu	Thr	Leu	Glu
	2!	515				2520				25	25				
Glu	Ala	Cys	Gln	Leu	Thr	Pro	Pro	His	Ser	Ala	Arg	Ser	ГÀЗ	Tyr	Gly
2	530				253	5			25	540					
Phe	Gly	Ala	Lys	Glu	Val	Arg	Ser	Leu	Ser	Gly	Arg	Ala	Val	Asn	His
2545	5			255	50			2	555				256	0	
Ile	Lys	Ser	Val	Trp	Lys	Asp	Leu	Leu	Glu	Asp	Pro	Gln	Thr	Pro	Ile
			256	55			2	570				257	5		
Pro	Thr	Thr	Ile	Met	Ala	Ŀуs	Asn	G1u	Val	Phe	Cys	Val	Asp	Pro	Ala
		25	80			;	2585				259	0			
Lys	Gly	Gly	Lys	Lys	Pro	Ala	Arg	Leu	Ile	Val	Tyr	Pro	Asp	Leu	Gly
	2 !	595				2600	t			26	05				
Val	Arg	Val	Cys	Glu	Lys	Met	Ala	Leu	Tyr	Asp	Ile	Thr	Gln	Lys	Leu
2	610				261	5			2.6	520					

Pro	Gln	Ala	Val	Met	Gly	Ala	Ser	Tyr	Gly	Phe	Gln	Tyr	Ser	Pro	Ala
2625	á			263	0			2	635				2640	+	
Gln	Arg	Va1	Glu	Tyr	Leu	Leu	Lys	Ala	Trp	Ala	G1 u	Ľуs	Lys	Asp	Pro
			264	5			2	650				2655	i		
Met	Gly	Phe	Ser	Tyr	Asp	Thr	Arg	Сув	Phe	Asp	Ser	Thr	Va1	Thr	Glu
		26	60			2	665				267	0			
Arg	Asp	Ile	Arg	Thr	Glu	Glu	Ser	Ile	Tyr	Gln	Ala	Суа	Ser	Leu	Pro
	26	75				2680				268	35				
Glu	Glu	Ala	Arg	Thr	Ala	Ile	His	Ser	Leu	Thr	Glu	Arg	Leu	Tyr	Val
2	690				269	5			27	00					
Gly	Gly	Pro	Met	Phe	Asn	Ser	Lys	Gly	Gln	Thr	Сув	Gly	Tyr	Arg	Arg
2709	5			271	0			2	715				2720	1	
Cys	Arg	Ala	Ser	Gly	Val	Leu	Thr	Thr	Ser	Met	Gly	Asn	Thr	Ile	Thr
			272	5			2	730				2735	;		
Cys	Tyr	Val	Lys	Ala	Leu	Ala	Ala	Сув	Lys	Ala	Ala	Gly	Ile	Val	Ala
		27	4 0			2	745				275	0			
Pro	Thr	Met	Leu	Val	Cys	Gly	Asp	Asp	Leu	Val	Val	Ίle	Ser	Glu	ser
	27	55				2760				276	55				
Gln	Gly	Thr	Glu	Glu	Asp	Glu	Arg	Asn	Leu	Arg	Ala	Phe	Thr	Glu	Ala
2	770				277	5			27	80					
Met	Thr	Arg	Tyr	Ser	Ala	Pro	Pro	Gly	Asp	Pro	Pro	Arg	Pro	Glu	Tyr
2785	5			279	0			2	795				2800	)	
Asp	Leu	Glu	Leu	Ile	Thr	Ser	Cys	ser	ser	Asn	Val	ser	Val	Ala	Leu
			280	5			2	810				2815	;		
Gly	Pro	Arg	Gly	Arg	Arg	Arg	Tyr	Tyr	Leu	Thr	Arg	Asp	Pro	Thr	Thr
		28	20			2	825				283	0			
Pro	Leu	Ala	Arg	Ala	Ala	Trp	Glu	Thr	Val	Arg	His	Ser	Pro	Ile	Asn
	28	35				2840				289	<del>1</del> 5				
Ser	Trp	Leu	Gly	Asn	Ile	Ile	Gln	Tyr	Ala	Pro	Thr	Ile	Trp	Val	Arg
2	850				285	5			28	60					
Met	Val	Leu	Met	Thr	His	Phe	Phe	Ser	Ile	Leu	Met	Val	Gln	Asp	Thr
2865	:			287	n			2	875				2880	1	

```
Leu Asp Gln Asn Leu Asn Phe Glu Met Tyr Gly Ser Val Tyr Ser Val
                             2890
            2885
                                               2895
Asn Pro Leu Asp Leu Pro Ala Ile Ile Glu Arg Leu His Gly Leu Asp
         2900
                          2905
                                            2910
Ala Phe Ser Met His Thr Tyr Ser His His Glu Leu Thr Arg Val Ala
                       2920
                                         2925
Ser Ala Leu Arg Lys Leu Gly Ala Pro Pro Leu Arg Val Trp Lys Ser
                                     2940
                    2935
Arg Ala Arg Ala Val Arg Ala Ser Leu Ile Ser Arg Gly Gly Lys Ala
                2950
                                  2955
                                                   2960
2945
Ala Val Cys Gly Arg Tyr Leu Phe Asn Trp Ala Val Lys Thr Lys Leu
                                               2975
            2965
                              2970
Lys Leu Thr Pro Leu Pro Glu Ala Arg Leu Leu Asp Leu Ser Ser Trp
                                            2990
         2980
                           2985
Phe Thr Val Gly Ala Gly Gly Gly Asp Ile Phe His Ser Val Ser Arg
      2995
                      3000
                                        3005
Ala Arg Pro Arg Ser Leu Leu Phe Gly Leu Leu Leu Phe Val Gly
  3010
                   3015
                                    3020
Val Gly Leu Phe Leu Leu Pro Ala Arg
3025
               3030
<210> 5
<211> 9674
```

<220>

<221> CDS

<212> DNA

<222> (341).,(9442)

<213> Hepatitis C virus

<400> 5

acco	gcc	cat a	aatag	agaad	g ac	cacto	cgcc	atg	jaato	act	ccc	etgte	gag	gaact	actgt	60
ctto	caege	cag a	aaago	egtet	a go	catç	gggt	tag	gtate	gagt	gte	jtaca	agc	ctcca	ıggccc	120
ccc	ecte	aag i	ggaga	agcca	at ag	gtggt	cetge	ggs	aaccg	ggtg	agta	caco	egg	aattg	lcc333	180
aaga	actg	ggt '	ccttl	tatto	gg at	caaa	ccac	tet	tatgo	eceg	gdda	atttg	<b>a</b> aa	cgtgo	ccccg	240
caaç	gact	gat :	agcc	gagta	ıg cg	gttgg	;gttg	cga	aaagg	ject	tgt	ggtad	etg	actga	ıtaggg	300
tgal	tge	gag	tgcc	ceggg	ja g	gtete	gtag	ace	gtgo	acc	atg	agc	aca	aat	ccc	355
									Me	t Se	r Th	r As	n Pa	0		
									,	i.			5			
															<b>a</b> 2.0	403
														caa		403
гла	Pro	GIn			Tnr	ьув			Thr	ASI	Arg		Pro	Gln	Авр	
			1	0				15				20				
att	aaq	ttt	ccq	qqc	ggc	qqc	cag	atc	gtt	gge	gga	gta	tac	ttg	ttg	451
														Leu		
	-		25	-			3 0				3 9					
ccg	cgc	agg	ggc	ccc	agg	ttg	ggt	gŧg	cgc	gcg	aca	agg	aag	gct	tcg	499
Pro	Arg	Arg	Gly	Pro	Arg	Leu	Gly	Val	Arg	Ala	Thr	Arg	Lys	Ala	Ser	
		40				45				5	0					
gag	cgg	tee	cag	cca	cgt	999	agg	cgc	cag	ccc	atc	ccc	aaa	cat	cgg	547
Glu	Arg	Ser	Gln	Pro	Arg	Gly	Arg	Arg	Gln	Pro	Ile	Pro	Lys	His	Arg	
	55				60				6	55						
														ccc		595

70				7	5				80				85			
tat	999	aat	gag	<b>3</b> 33	ctc	ggt	tgg	gca	gga	tgg	ctc	ctg	tcc	aet	cga	643
Tyr	Gly	Asn	Glu	Gly	Leu	Gly	Trp	Ala	Gly	Trp	Leu	Leu	ser	Pro	Arg	
			9	0				95				100				
ggt	tcc	cgt	ccc	tca	tgg	ggc	ccc	aat	gac	ccc	cgg	cat	agg	tcg	cgc	691
Gly	Ser	Arg	Pro	Ser	Trp	Gly	Pro	Asn	Asp	Pro	Arg	His	Arg	Ser	Arg	
		1	05				110				11	5				
															ctc	739
Asn	Val	Gly	Lys	Val	Ile	Asp	Thr	Leu	Thr	Сув	Gly	Phe	Ala	Asp	Leu	
	1	20				125				13	0					
															agt	787
Leu	Gly	Tyr	Val	Pro	Val	Val	Gly	Ala	Pro	Leu	Ser	Gly	Val	Ala	Ser	
	135				14(	)			1	45						
															gca	835
	Leu	Ala	His	-		Arg	Val		Glu	Asp	Gly	Val		Phe	Ala	
150				15	5			-	1.60				165			
															cta -	883
Thr	Gly	Asn			Gly	Cys			Ser	Ile	Phe			Ala	Leu	
			17	0				175				180				
ata	tido	tac	atc	act	act	cca	atc	tet	get	ate	caa	ata	aao	aac	acc	931
_		-				_	•		Ala	_			-			
		-	85				190				19		•			
		~					-									
agc	aac	gcc	tat	atg	geg	act	aac	gac	tgt	tcc	aat	gac	agc	atc	act	979

Ser Asn Ala Tyr Met Ala Thr Asn Asp Cys Ser Asn Asp Ser Ile Thr

	2	00				205				21	. 0					
		att	~~~	gcc	<i>a</i> aa	ata	ata	ast	ata	000	aaa	tac	ata	cca	tac	1027
				Ala												
		ьец	GIU	Ala	22(		ьец	ura		25	GIY	Cyn	Vai	210	cys	
	215				22(	,			2	23						
gag	aaa	atg	<b>ggg</b>	aac	aca	tca	cgg	tgc	tgg	ata	cca	gtc	tca	cca	aac	1075
				Asn												
230				23	5			2	240				245			
gtg	gct	gtg	cgg	cag	ect	ggc	gc¢	ctc	acg	cgg	ggc	ttg	cgg	acg	cac	1123
Val	Ala	۷al	Arg	Gln	Pro	Gly	Ala	Leu	Thr	Arg	Gly	Leu	Arg	Thr	His	
			25	0			:	255				260				
atc	gac	atg	gtc	gtg	ttg	tcc	gcc	acg	ctc	tgc	tcc	gct	ctc	tac	gtg	1171
Ile	Asp	Met	Val	Val	Leu	Ser	Ala	Thr	Leu	Сув	Ser	Ala	Leu	Tyr	Val	
		2	65				270				27	5				
ggg	gac	ctc	tgt	ggc	999	gtg	atg	ctc	gcg	tcc	cag	atg	ttc	att	gtc	1219
Gly	Asp	Leu	Сув	Gly	Gly	Val	Met	Leu	Ala	Ser	Gln	Met	Phe	Ile	Val	
	2	80				285				29	0 0					
tcg	ccg	cag	cac	cac	tgg	ttc	gtg	cag	gaa	tgc	aat	tgc	tcc	atc	tac	1267
Ser	Pro	Gln	His	His	Trp	Phe	Val	Gln	Glu	Сув	Asn	Cys	Ser	Ile	Tyr	
	295				300	)			3	05						
cct	ggc	gcc	atc	act	999	cac	cgt	atg	gca	tgg	gac	atg	atg	atg	aac	1315
Pro	Gly	Ala	Ile	Thr	Gly	His	Arg	Met	Ala	Trp	Asp	Met	Met	Met	Asn	
310				31	5			:	320				325			
tgg	tcg	ccc	acg	acc	acc	atg	atc	ctg	gcg	tac	gtg	atg	cgc	gtt	ccc	1363

Trp Ser Pro Thr Thr Met Ile Leu Ala Tyr Val Met Arg Val Pro

gag	gtc	atc	ata	gac	atc	att	agc	gga	gct	cac	tgg	ggc	gtc	atg	ttt	1411
Glu	Val	Ile	Ile	Asp	Ile	Ile	Ser	Gly	Ala	His	Trp	Gly	Val	Met	Phe	
		3	45				350				35	5				
gge	ctg	gcc	tac	ttc	tct	atg	cag	gga	gcg	tgg	gcg	aag	gtc	gtt	gtc	1459
Gly	Leu	Ala	Tyr	Phe	Ser	Met	Gln	Gly	Ala	Trp	Ala	Lys	Val	Val	Val	
	3	360				365				3 7	0					
atc	ctc	ctg	ctg	gcc	tat	aaa	gtg	gac	gcg	tac	acc	acc	acg	act	999	1507
Ile	Leu	Leu	Leu	Ala	Ser	Gly	Val	Asp	Ala	Tyr	Thr	Thr	Thr	Thr	Gly	
	375				380	}			3	85						
agc	gct	gct	aaa	cgc	act	acc	agt	agc	ctg	gcc	agc	gcc	ttc	tcc	cct	1555
Ser	Ala	Ala	Gly	Arg	Thr	Thr	Ser	Ser	Leu	Ala	Ser	Ala	Phe	Ser	Pro	
390				3 9	5			4	00				405			
ggc	gct	cgg	cag	aac	att	cag	ctc	att	aat	acc	aat	ggt	agc	tgg	cac	1603
Gly	Ala	Arg	Gln	Asn	Ile	Gln	Leu	Ile	Asn	Thr	Asn	Gly	Ser	Trp	His	
			41	0			•	115				420				
atc																
<b>-1</b> -	aac	cgc	acc	gcc	ctg	aat	tgc	aac	gat	tcc	ttg	cac	acc	ggc	ttc	1651
rre							-							ggc Gly		1651
rre		Arg					-					His				1651
ire		Arg	Thr				Сув				Leu	His				1651
	Asn	Arg	Thr 25	Ala	Leu	Asn	Cys 430	Asn	Asp	Ser	Leu 43!	His	Thr		Phe	1699
ttc	Asn acg	Arg 4: gcc	Thr 25 ctg	Ala	Leu tac	Asn atc	Cys 430 cat	Asn aag	Asp ttc	Ser aac	Leu 43! tcg	His 5	Thr gga	Gly	Phe ccc	
ttc	Asn acg Thr	Arg 4: gcc	Thr 25 ctg	Ala	Leu tac	Asn atc	Cys 430 cat	Asn aag	Asp ttc	Ser aac	Leu 43! tcg Ser	His 5	Thr gga	Gly tgt	Phe ccc	
ttc	Asn acg Thr	Arg 4: gcc Ala	Thr 25 ctg	Ala	Leu tac	Asn atc	Cys 430 cat	Asn aag	Asp ttc	Ser aac Asn	Leu 43! tcg Ser	His 5	Thr gga	Gly tgt	Phe ccc	

Glu Arg Leu Ser Ala Cys Arg Asn Ile Glu Asp Phe Arg Ile Gly Trp

			;													
ggc	gcc	ctg	caa	tac	gac	gac	aat	gtc	acc	aat	cca	gaa	gat	atg	agg	1795
Gly	Ala	Leu	Gln	Tyr	qaA	Asp	Asn	Val	Thr	Asn	Pro	Glu	Asp	Met	Arg	
470				47	5			4	180				485			
cca	tat	tgc	tgg	cac	tac	cca	cca	aaa	cag	tgt	ggc	gta	gtc	ccc	gca	1843
Pro	Tyr	Сув	Trp	His	Tyr	Pro	Pro	Lys	Gln	Сув	Gly	Val	Val	Pro	Ala	
			4 9	0				195				500				
aaa	acc	gtg	tgc	ggc	cca	gtg	tac	tgt	ttc	acc	cot	agc	ccg	gtg	gta	1891
Gly	Thr	Val	Cys	Gly	Pro	Val	Tyr	Cys	Phe	Thr	Pro	Ser	Pro	Val	Val	
		5	05				510				51	5				
gtg	ggc	acg	acc	gat	aga	ctt	gga	gtg	cct	act	tac	acg	tgg	gga	gag	1939
Va1	Gly	Thr	Thr	Asp	Arg	Leu	Gly	Val	Pro	Thr	Tyr	Thr	Trp	Gly	Glu	
	9	20				525				53	0					
							•									
aat	gag	aca	gat	gtc	ttc	cta	ttg	aac	agc	acc	cga	cca	ccg	tcg	999	1987
Asn	Glu	Thr	Asp	Val	Phe	Leu	Leu	Asn	Ser	Thr	Arg	Pro	Pro	Ser	Gly	
	535				540	)			5	4 5						
tca	tgg	ttt	ggc	tgc	acg	tgg	atg	aac	tcc	act	ggc	ttc	acc	aag	acc	2035
Ser	Trp	Phe	G1y	Сув	Thr	Trp	Met	Asn	Ser	Thr	Gly	Phe	Thr	Lys	Thr	
550				5 5	5			Ĩ	560				565			
tgc	ggc	gca	cca	ccc	tge	cgc	act	aga	gct	gac	ttc	aat	acc	agc	aca	2083
Cys	Gly	Ala	Pro	Pro	Cys	Arg	Thr	Arg	Ala	Asp	Phe	Asn	Thr	Ser	Thr	
			57	0			!	575				580				
		h h =	baa		200	~~~	+~+	F 4- 4-	200	222	as t	cet	a==	acc	act	2131

Asp Leu Leu Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu Ala Thr

tac	atc	aaa	tgt	ggt	tcc	aaa	cct	tgg	ctc	acg	cca	aag	tgt	ctg	gtt	2179
Tyr	Ile	Lys	Суз	Gly	Ser	Gly	Pro	Trp	Leu	Thr	Pro	Lys	Cys	Leu	Val	
	(	500				605				61	. 0					
gac	tac	ccc	tac	agg	ctc	tgg	cat	tac	cct	tgc	aca	gtc	aat	tac	tcc	2227
Asp	Tyr	Pro	Tyr	Arg	Leu	Trp	His	Tyr	Pro	Cys	Thr	Val	Asn	Tyr	Ser	
	615				620	)			6	25						
acc	ttc	aag	atc	agg	atg	tat	gtg	aaa	gga	gtt	gag	cac	agg	ctc	atg	2275
Thr	Phe	Lys	Ile	Arg	Met	Tyr	Val	Gly	Gly	Va1	Glu	His	Arg	Leu	Met	
630				63	5			6	540				645			
gcc	g¢g	tgc	aat	ttc	act	cgt	999	gat	cgc	tgc	aac	ttg	gag	gat	agg	2323
Ala	Ala	Cys	Asn	Phe	Thr	Arg	Gly	Asp	Arg	Сув	Asn	Leu	Glu	Asp	Arg	
			65	0				655				660				
gac	aga	agt	caa	cag	act	cct	ctg	ttg	cac	tcc	acc	acg	gaa	tgg	gcc	2371
Asp	Arg	Ser	Gln	Gln	Thr	Pro	Leu	Leu	His	Ser	Thr	Thr	Glu	Trp	Ala	
		6	65				670				67	5				
att	ttg	ccc	tgc	tct	ttc	tca	gac	ttg	acc	gct	ttg	tcg	act	ggt	ctt	2419
Ile	Leu	pro	Сув	Ser	Phe	Ser	Asp	Leu	Pro	Ala	Leu	Ser	Thr	Gly	Leu	
	6	80				685				69	0					
ata	cac	ctc	cac	caa	aat	atc	gtg	gac	gta	caa	tat	atg	tat	ggc	ctg	2467
Leu	His	Leu	His	Gln	Asn	Ile	Va1	Asp	Val	Gln	Tyr	Met	Tyr	Gly	Leu	
	695				700	)			7	0 5		-				
tca	cct	qcc	ata	aca	caa	tat	atc	gtt	cga	tgg	gag	tgg	gta	gta	ctc	2515

Ser Pro Ala Leu Thr Gln Tyr Ile Val Arg Trp Glu Trp Val Val Leu

710				71	5			•	720				725			
tta	ttc	ctg	ctc	cta	gcg	gac	gcc	agg	gtc	tgc	gcc	tgc	ttg	tgg	atg	2563
Leu	Phe	Leu	Leu	Leu	Ala	Asp	Ala	Arg	Val	CAa	Ala	Cys	Leu	Trp	Met	
			73	0				735				740				
ctc	atc	ttg	ctg	ggc	caa	gcc	gaa	gca	gca	ctg	gag	aag	ctg	gtc	gtc	2611
Leu	Ile	Leu	Leu	Gly	Gln	Ala	Glu	Ala	Ala	Leu	Glu	Lys	Leu	Val	Val	
		7.	45				750				75!	5				
ttg	cac	gct	gcg	agc	gca	gct	agc	tgc	aat	ggc	ttc	ctg	tat	ttt	gtc	2659
Leu	His	Ala	Ala	ser	Ala	Ala	ser	Сув	Asn	G1y	Phe	Leu	Tyr	Phe	Val	
	7	760				765				77	0					
atc	ttt	ctc	ata	act	act	taa	cac	atc	aaq	aat	agg	ata	atc	ccc	ttg	2707
		Leu														
	775				780					85						
				att	act	aac	cta	taa	ccq	tte	tac	cta	eta	ata	cta	2766
gat	gct	tat	tcc	CCC	400	224	ceg	- 33	-		-3-		5	CCC		2755
		tat Tyr														275.
					Thr			Trp								2753
Ala 790	Ala	Туг	Ser	Leu 79	Thr 5	Gly	Leu	Trp	Pro	Phe	Сув	Leu	Leu 805	Leu		2803
Ala 790 gca	Ala ctg	Туг	Ser cag	Leu 79 cag	Thr 5 gct	Gly tac	Leu gcc	Trp	Pro 300 gat	Phe gca	Cys tct	Leu gtg	Leu 805 cac	Leu gga	Leu cag	
Ala 790 gca	Ala ctg	Tyr ccc	Ser cag	Leu 79 cag Gln	Thr 5 gct	Gly tac	Leu gcc Ala	Trp	Pro 300 gat	Phe gca	Cys tct	Leu gtg	Leu 805 cac	Leu gga	Leu cag	
Ala 790 gca Ala	Ala ctg Leu	ccc ¢ Pro	cag Gln 81	Leu 79 cag Gln 0	Thr 5 gct Ala	Gly tac Tyr	Leu gcc Ala	Trp tat Tyr 815	Pro 300 gat Asp	Phe gca Ala	Cys tct Ser	gtg Val 820	Leu 805 cac His	Leu gga Gly	Leu cag	
Ala 790 gca Ala gtg	Ala ctg Leu	ccc ¢ Pro	cag Gln 81 gct	Leu 79 cag Gln 0	Thr 5 gct Ala cta	Gly tac Tyr	gcc Ala	Trp tat Tyr 815 att	Pro 300 gat Asp	Phe gca Ala	tct ser	gtg Val 820	Leu 805 cac His	Leu gga Gly	Leu cag Gln	2803
Ala 790 gca Ala gtg	Ala ctg Leu	ccc Pro	cag Gln 81 gct	Leu 79 cag Gln 0	Thr 5 gct Ala cta	Gly tac Tyr	gcc Ala	Trp tat Tyr 815 att	Pro 300 gat Asp	Phe gca Ala	tct ser	gtg Val 820 aca	Leu 805 cac His	Leu gga Gly	Leu cag Gln	2803
Ala 790 gca Ala gtg	Ala ctg Leu ggc	ccc Pro gcg Ala 8:	cag Gln 81 gct Ala	Leu 79 cag Gln 0 ttg Leu	Thr 5 gct Ala cta Leu	tac Tyr gta Val	gcc Ala ctg Leu 830	tat Tyr 815 att	Pro 300 gat Asp acc	gca Ala ctc Leu	tct Ser ttt Phe	gtg Val 820 aca Thr	Leu 805 cac His ctc	gga Gly acc	cag Gln ccg Pro	2803 2851
Ala 790 gca Ala gtg Val	Ala ctg Leu ggc Gly	ccc Pro	cag Gln 81 gct Ala 25	Leu 79 cag Gln 0 ttg Leu	Thr 5 gct Ala cta Leu	Gly tac Tyr gta Val	gcc Ala ctg Leu 830	tat Tyr 315 att Ile	Pro 300 gat Asp acc Thr	gca Ala ctc Leu	tct Ser ttt Phe 839	gtg Val 820 aca Thr	Leu 805 cac His ctc Leu	gga Gly acc Thr	cag Gln ccg Pro	2803

ctg	acc	ctg	gcg	gaa	acc	atg	gtc	cag	gag	tgg	gca	cca	tcc	atg	cag	2947
Leu	Thr	Leu	Ala	Glu	Thr	Met	Val	Gln	Glu	Trp	Ala	Pro	Ser	Met	Gln	
8	855				860	)			8	65						
gcg	cgc	ggc	ggc	cgŧ	gat	ggc	atc	ata	tgg	gcc	gcc	acc	ata	ttt	tgc	2995
Ala	Arg	Gly	Gly	Arg	Asp	Gly	Ile	Ile	Trp	Ala	Ala	Thr	Ile	Phe	Cys	
870				87	5			8	380				885			
													,			
ccg	ggc	gta	gtg	ttt	gac	ata	acc	aag	tgg	ctc	tta	gcg	gtg	ctt	aaa	3043
Pro	Gly	Val	Val	Phe	Asp	Ile	Thr	гуз	Trp	Leu	Leu	Ala	Va1	Leu	Gly	
			89	0				895				900				
cct	ggt	tac	ctc	cta	aga	ggt	gct	ttg	acg	cgc	gtg	cca	tat	ttc	gtc	3091
Pro	Gly	Tyr	Leu	Leu	Arg	Gly	Ala	Leu	Thr	Arg	Val	Pro	тух	Phe	Val	
		9	05				910				91	5				
		9	05				910				91	5				
aga	gcc			ctg	ctg	aga		tgc	act	atg			cac	ata	gcg	3139
aga Arg		cac	gct				atg				gtg	agg				3139
	Ala	cac	gct				atg				gtg Val	agg				3139
	Ala	cac His	gct			Arg	atg			Met	gtg Val	agg				3139
	Ala 9	cac His	gct Ala	Leu	Leu	Arg 925	atg Met	Сув	Thr	Met 93	gtg Val	agg Arg	His	Leu	Ala	3139
Arg	Ala 9	cac His 20 agg	gct Ala	Leu gtc	Leu	Arg 925 atg	atg Met gcg	Cys cta	Thr	Met 93 gcc	gtg Val 0	agg Arg ggc	His agg	Leu tgg	Ala	
ggly gga Arg	Ala 9	cac His 20 agg	gct Ala	Leu gtc	Leu	Arg 925 atg Met	atg Met gcg	Cys cta	Thr tta Leu	Met 93 gcc	gtg Val 0	agg Arg ggc	His agg	Leu tgg	Ala	
ggly gga Arg	Ala 9 ggt Gly	cac His 20 agg	gct Ala	Leu gtc	Leu cag Gln	Arg 925 atg Met	atg Met gcg	Cys cta	Thr tta Leu	Met 93 gcc Ala	gtg Val 0	agg Arg ggc	His agg	Leu tgg	Ala	
Arg ggg Gly	Ala ggt Gly	cac His 20 agg Arg	gct Ala tac Tyr	Leu gtc Val	cag Gln 940	Arg 925 atg Met	atg Met gcg Ala	Cys cta Leu	Thr tta Leu 9	93 gcc Ala 45	gtg Val 0 ctt Leu	agg Arg ggc Gly	His agg Arg	Leu tgg Trp	Ala	3187
Arg ggg Gly	Ala ggt Gly 935 act	cac His 20 agg Arg	gct Ala tac Tyr	Deu gtc Val	cag Gln 940	Arg 925 atg Met	atg Met gcg Ala	Cys cta Leu	tta Leu 9	Met 93 gcc Ala 45	gtg Val 0 ctt Leu	agg Arg ggc Gly	His agg Arg	tgg Trp	Ala act Thr	3187
agg ggc	Ala ggt Gly 935 act	cac His 20 agg Arg	gct Ala tac Tyr	Deu gtc Val	cag Gln 940 gac Asp	Arg 925 atg Met	atg Met gcg Ala	Cys cta Leu acc	tta Leu 9	Met 93 gcc Ala 45	gtg Val 0 ctt Leu	agg Arg ggc Gly	His agg Arg	tgg Trp gct	Ala act Thr	3187
agg gly	Ala ggt Gly 935 act	cac His 20 agg Arg	gct Ala tac Tyr	gtc Val tat	cag Gln 940 gac Asp	Arg 925 atg Met	atg Met gcg Ala	Cys cta Leu acc	tta Leu 9 cct	Met 93 gcc Ala 45	gtg Val 0 ctt Leu	agg Arg ggc Gly	agg Arg tgg	tgg Trp gct	Ala act Thr	3187

Ser Gly Leu Arg Asp Leu Ala Val Ala Val Glu Pro Ile Ile Phe Ser

				•												
ccg	atg	gag	aag	aaa	gtc	atc	gtt	tgg	gga	gcg	gag	acg	gct	gcg	tgc	3331
Pro	Met	Glu	Lys	Lys	Val	Ile	Val	Trp	Gly	Ala	Glu	Thr	Ala	Ala	СЛа	
		9	85				990				99	5				
999	gac	atc	ttg	cac	gga	ctt	acc	gtg	tcc	gcc	cga	ctc	ggt	cgg	gag	3379
Gly	Asp	Ile	Leu	His	Gly	Leu	Pro	Val	Ser	Ala	Arg	Leu	Gly	Arg	Glu	
	10	000				1005				10	10					
atc	ata	ctt	ggc	cca	gct	gat	ggc	tac	acc	tcc	aag	aaa	tgg	aag	ctt	3427
Ile	Leu	Leu	Gly	Pro	Ala	Asp	Gly	Tyr	Thr	Ser	Lys	Gly	Trp	Lys	Leu	
1	015				102	0			10	25						
ctc	gcc	ccc	atc	acc	gct	tac	gcc	cag	cag	aca	cga	ggt	ctc	ttg	ggc	3475
Leu	Ala	Pro	Ile	Thr	Ala	Tyr	Ala	Gln	Gln	Thr	Arg	Gly	Leu	Leu	Gly	
1030	)			103	5			1	040				104	5		
tet	ata	gtg	gtg	agc	a't g	acg	999	cgt	gac	aag	aca	gaa	cag	gcc	999	3523
ser	Ile	Val	Val	Ser	Met	Thr	Gly	Arg	Asp	ьув	Thr	Glu	Gln	Ala	Gly	
			105	0			1	055				1060	0			
gag	gtc	caa	gtc	ctg	tcc	aca	gtc	act	cag	tcc	ttc	ctc	gga	aca	tee	3571
Glu	Val	Gln	Val	Leu	Ser	Thr	Val	Thr	Gln	Ser	Phe	Leu	Gly	Thr	Ser	
		10	65			:	1070				107	5				
att	tcg	<b>a</b> aa	gtc	tta	tgg	act	gtt	tac	cac	gga	gct	ggc	aac	aag	aca	3619
Ile	Ser	Gly	Val	Leu	Trp	Thr	Val	Tyr	His	Gly	Ala	Gly	Asn	Lys	Thr	
	10	80				1085				10	90					
cta	gcc	ggc	tcg	cgg	ggc	ccg	gtc	acg	cag	atg	tac	tcg	agc	gcc	gag	3667

Leu Ala Gly Ser Arg Gly Pro Val Thr Gln Met Tyr Ser Ser Ala Glu

1095	11	00	1105		
			ect cct ggg a		
1110	1115		1120	1125	
1110	1113		1120	11.00	•
ccg tgt acg	tgt gga gc	g gte gae e	tg tat ttg g	gte aeg egg	aac gct 3763
			eu Tyr Leu V		
-	1130	113		1140	
gat gtc atc	ccg gct cg	a aga cgc g	ıgg gac aag d	cgg gga gcg	ctg ctc 3811
Asp Val Ile	Pro Ala Ar	g Arg Arg G	ly Asp Lys A	Arg Gly Ala	Leu Leu
11-	45	1150	•	1155	
tee eeg aga	ccc ctt tc	g acc ttg a	ag ggg tee t	tog ggg gga	cct gtg 3859
Ser Pro Arg	Pro Leu Se	r Thr Leu L	ys Gly Ser S	Ser Gly Gly	Pro Val
1160		1165	117	0	
ctt tgc cct	agg ggc ca	c got gto g	ga atc ttc	cgg gca gct	gtg tgc 3907
Leu Cys Pro	Arg Gly Hi	s Ala Val G	Sly Ile Phe A	Arg Ala Ala	Val Cys
1175	11	80	1185		
tct cgg ggt	gtg gct aa	g tee ata g	jat ttc atc o	ccc gtt gag	acg ctc 3955
Ser Arg Gly	Val Ala Ly	s Ser Ile A	asp Phe Ile i	Pro Val Glu	Thr Leu
1190	1195		1200	1205	i .
gae ate gte	acg egg to	t ccc acc t	tt agt gac a	aac agc aca	cca cca 4003
Asp Ile Val	Thr Arg Se	r Pro Thr P	he Ser Asp <i>l</i>	Asn Ser Thr	Pro Pro
	1210	121	15	1220	
get gtg ccc	cag acc ta	t cag gtg g	gg tac ttg	cac gcc ccc	act ggc 4051

Ala Val Pro Gln Thr Tyr Gln Val Gly Tyr Leu His Ala Pro Thr Gly

agt gga aaa	agc acc	aag	gtc	ccc	gtc	gcg	tac	gcc	gcc	cag	<b>3</b> 33	tat	4099
Ser Gly Lys	Ser Thr	Lys	Val	Pro	Val	Ala	Tyr	Ala	Ala	Gln	Gly	Tyr	
1240			1245				12	50					
	,												
aaa gtg ctg	gtg ctc	aat	ccc	tcg	gtg	gct	gcc	acc	ctg	gga	ttt	999	4147
Lys Val Leu	Val Leu	Asn	Pro	Ser	Val	Ala	Ala	Thr	Leu	Gly	Phe	Gly	
1255		1260	0			12	65						
geg tae ttg	tcc aag	gca	cat	ggc	atc	aac	ccc	aac	att	agg	act	gga	4195
Ala Tyr Leu	Ser Lys	Ala	His	Gly	Ile	Asn	Pro	Asn	Ile	Arg	Thr	Gly	
1270	12	75			1	280				128	5		
gtc aga act	gtg acg	acc	ggg	gag	aac	att	aca	tac	tcc	acg	tat	ggt	4243
Val Arg Thr	Val Thr	Thr	Gly	Glu	Pro	Ile	Thr	Tyr	Ser	Thr	Tyr	Gly	
	1290				295				1300	)			
					295				1300	)			
aaa ttc ctc	1290			1		ggc	ggc	gcc			atc	atc	4291
aaa ttc ctc Lys Phe Leu	1290 gcc gat	<b>3</b> 33	ggc	1 tgc	gca				tat	gac			4291
Lys Phe Leu	1290 gcc gat	<b>3</b> 33	ggc Gly	1 tgc	gca				tat Tyr	gac			4291
Lys Phe Leu	1290 gcc gat Ala Asp	<b>3</b> 33	ggc Gly	1 tgc Cys	gca			Ala	tat Tyr	gac			4291
Lys Phe Leu	1290 gcc gat Ala Asp 05	д <b>1</b> у	ggc Gly	tgc Cys	gca Ala	Gly	Gly	Ala 131	tat Tyr 5	gac Asp	Ile	Ile	4291 4339
Lys Phe Leu 13	1290 gcc gat Ala Asp 05 gaa tgc	ggg Gly cac	ggc Gly 1	tgc Cys 310	gca Ala gat	Gly gct	Gly	Ala 131 act	tat Tyr 5	gac Asp	Ile ggc	Ile atc	
Lys Phe Leu 13 ata tgc gat	1290 gcc gat Ala Asp 05 gaa tgc	ggg Gly cac	ggc Gly 1	tgc Cys 310 gtg Val	gca Ala gat	Gly gct	Gly	Ala 131 act Thr	tat Tyr 5	gac Asp	Ile ggc	Ile atc	
Lys Phe Leu 13 ata tgc gat Ile Cys Asp	1290 gcc gat Ala Asp 05 gaa tgc	ggg Gly cac	ggc Gly 1 tct Ser	tgc Cys 310 gtg Val	gca Ala gat	Gly gct	Gly acc Thr	Ala 131 act Thr	tat Tyr 5	gac Asp	Ile ggc	Ile atc	
Lys Phe Leu 13 ata tgc gat Ile Cys Asp	1290 gcc gat Ala Asp 05 gaa tgc Glu Cys	aga Gly His	ggc Gly tct Ser 1325	tgc Cys 310 gtg Val	gca Ala gat Asp	Gly gct Ala	Gly acc Thr	Ala 131 act Thr	tat Tyr 5 att Ile	gac Asp ctc Leu	Ile ggc Gly	Ile atc Ile	4339
Lys Phe Leu 13 ata tgc gat Ile Cys Asp 1320	1290 gcc gat Ala Asp 05 gaa tgc Glu Cys	ggg Gly cac His	ggc Gly 1 tct Ser 1325 gca	tgc Cys 310 gtg Val	gca Ala gat Asp	Gly gct Ala gcc	acc Thr 13:	Ala 131 act Thr 30	tat Tyr 5 att Ile	gac Asp ctc Leu	Ile ggc Gly	Ile atc Ile	4339
Lys Phe Leu 13 ata tgc gat Ile Cys Asp 1320 ggg aca gtc	1290 gcc gat Ala Asp 05 gaa tgc Glu Cys	ggg Gly cac His	ggc Gly tct Ser 1325 gca Ala	tgc Cys 310 gtg Val	gca Ala gat Asp	gct Ala gcc Ala	acc Thr 13:	Ala 131 act Thr 30	tat Tyr 5 att Ile	gac Asp ctc Leu	Ile ggc Gly	Ile atc Ile	4339
Lys Phe Leu 13  ata tgc gat Ile Cys Asp 1320  ggg aca gtc Gly Thr Val	1290 gcc gat Ala Asp 05 gaa tgc Glu Cys ctt gac Leu Asp	ggg Gly cac His caa	ggc Gly tct Ser 1325 gca Ala	tgc Cys 310 gtg Val	gca Ala gat Asp	gct Ala gcc Ala	acc Thr 13:	Ala 131 act Thr 30	tat Tyr 5 att Ile	gac Asp ctc Leu	Ile ggc Gly	Ile atc Ile	4339

Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Thr Pro His Pro Asn

1350 1355				1360					1365						
ata gag	gag	gta	gcc	ata	gga	cag	gag	ggt	gag	atc	ccc	ttc	tat	ggg	4483
Ile Glu	Glu	Val	Ala	Leu	Gly	Gln	Glu	Gly	Glu	Ile	Pro	Phe	Tyr	Gly	
		137	0			1	375				1380	)			
agg gcg	ttt	ccc	ctg	tet	tac	atc	aag	gga	ggg	agg	cac	ttg	att	tta	4531
Arg Ala	Phe	Pro	Leu	ser	Tyr	Ile	Lys	Gly	Gly	Arg	His	Leu	Ile	Phe	
	13	85			1	1390				139	5				
tgc cac	tca	aag	aaa	aag	tgt	gac	gag	ata	gca	acg	gcc	ctt	cgg	ggc	4579
Cys His	Ser	Lys	Lys	Lys	Cys	Asp	Glu	Leu	Ala	Thr	Ala	Leu	Arg	Gly	
1	400				1405				14	10					
atg ggc	ttg	aac	gat	gtg	gca	tat	tac	aga	999	ttg	gac	gtc	tcc	ata	4627
Met Gly	Leu	Asn	Ala	Val	Ala	Tyr	Tyr	Arg	Gly	Leu	Asp	Val	Ser	Ile	
1415				142	0			14	25						
1415				142	0			14	25						
1415 ata cca	act	caa	gga			gtg	gtc			acc	gac	gac	ata	atg	4675
				gat	gtg			gtt	gcc						4675
ata cca				gat Asp	gtg		Va1	gtt	gcc				Leu		4675
ata cca Ile Pro			Gly	gat Asp	gtg		Va1	gtt Val	gcc			Ala	Leu		4675
ata cca Ile Pro	Thr	Gln	Gly 143	gat Asp 5	gtg Val	Val	Val	gtt Val 440	gcc Ala	Thr	Asp	Ala 1449	Leu ;	Met	4675 4723
ata cca Ile Pro 1430	Thr	Gln act	Gly 143 gga	gat Asp 5 gac	gtg Val	Val gac	Val 1	gtt Val 440 gtg	gcc Ala	Thr	Asp tgc	Ala 1449 aac	Leu ; gta	Met gcg	
ata cca Ile Pro 1430	Thr	Gln act	Gly 143 gga Gly	gat Asp 5 gac	gtg Val	Val gac Asp	Val 1	gtt Val 440 gtg	gcc Ala	Thr	Asp tgc	Ala 1449 aac Asn	Leu ; gta	Met gcg	
ata cca Ile Pro 1430	Thr	Gln act Thr	Gly 143 gga Gly	gat Asp 5 gac	gtg Val	Val gac Asp	Val 1 tcc Ser	gtt Val 440 gtg	gcc Ala	Thr	Asp tgc Cys	Ala 1449 aac Asn	Leu ; gta	Met gcg	
ata cca Ile Pro 1430	Thr tat Tyr	act Thr	Gly 143 gga Gly	gat Asp 5 gac Asp	gtg Val ttt Phe	Val gac Asp	Val tcc Ser 455	gtt Val 440 gtg Val	gcc Ala atc	Thr gac Asp	tgc Cys	Ala 1445 aac Asn	Leu ; gta Val	gcg Ala	4723
ata cca Ile Pro 1430 acg ggg Thr Gly	Thr tat Tyr	Gln act Thr 145	Gly 143 gga Gly 0	gat Asp 5 gac Asp	gtg Val ttt Phe	yal gac Asp 1	Val tcc Ser 455	gtt Val 440 gtg Val	gcc Ala atc Ile	Thr gac Asp	tgc Cys 1460	Ala 1449 aac Asn	Leu gta Val ata	gcg Ala	4723
ata cca Ile Pro 1430  acg ggg Thr Gly	Thr tat Tyr	act Thr 145	Gly 143 gga Gly 0	gat Asp 5 gac Asp	gtg Val ttt Phe	yal gac Asp 1	Val tcc Ser 455	gtt Val 440 gtg Val	gcc Ala atc Ile	Thr gac Asp	tgc Cys 1460 ttc	Ala 1449 aac Asn	Leu gta Val ata	gcg Ala	4723
ata cca Ile Pro 1430  acg ggg Thr Gly	Thr tat Tyr cag	act Thr 145	Gly 143 gga Gly 0	gat Asp 5 gac Asp	gtg Val ttt Phe	yac Asp 1 agc	Val tcc Ser 455	gtt Val 440 gtg Val	gcc Ala atc Ile	Thr gac Asp acc	tgc Cys 1460 ttc	Ala 1449 aac Asn	Leu gta Val ata	gcg Ala	4723
ata cca Ile Pro 1430  acg ggg Thr Gly	tat Tyr cag Gln	act Thr 145 gcc Ala	Gly 143 gga Gly 0 gta Val	gat Asp gac Asp gac	gtg Val ttt Phe ttc	gac Asp 1 agc Ser	Val tcc Ser 455 ctg	gtt Val 440 gtg Val gac Asp	gcc Ala atc Ile	gac Asp acc Thr	tgc Cys 1460 ttc Phe	Ala  1449  aac  Asn  act  Thr	Leu gta Val ata Ile	gcg Ala acc	4723
ata cca Ile Pro 1430  acg ggg Thr Gly  gte acc Val Thr	tat Tyr cag Gln 14	act Thr 145 gcc Ala 65	Gly 143 gga Gly 0 gta Val	gat Asp 5 gac Asp caa	gtg Val ttt Phe	gac Asp 1 agc Ser 470	tcc ser 455 ctg Leu	gtt Val 440 gtg Val gac Asp	gcc Ala atc Ile ccc Pro	gac Asp acc Thr 147	tgc Cys 1460 ttc Phe 5	Ala  1449  aac  Asn  act  Thr	Leu gta Val ata Ile	gcg Ala acc Thr	4723 4771

1480	1485	1490	
ege aeg ggt aga g Arg Thr Gly Arg G			
1495	1500	1505	
gag cga gcc tca g	gga atg ttt gac	agt gta gta ctc	tgt gag tgc tac 4915
Glu Arg Ala Ser G	Gly Met Phe Asp	Ser Val Val Leu	Cys Glu Cys Tyr
1510	1515	1520	1525
gac gca gga gct g	got tgg tat gag	ctc tca cca gtg	gag acg acc gtc 4963
Asp Ala Gly Ala A	Ala Trp Tyr Glu	Leu Ser Pro Val	Glu Thr Thr Val
1530	) 1	535	1540
agg ctc agg gcg t	tat ttc aac acg	cet gge ttg cet	gtg tgc cag gac 5011
Arg Leu Arg Ala T	Tyr Phe Asn Thr	Pro Gly Leu Pro	Val Cys Gln Asp
1545	1550	155	5
cac ctt gag ttt t	tgg gag gca gtt	ttc acc ggc ctc	aca cac ata gac 5059
His Leu Glu Phe T	Irp Glu Ala Val	Phe Thr Gly Leu	Thr His Ile Asp
1560	1565	1570	
get cat tte ett t	ccc cag aca aag	cag tcg ggg gaa	aat ttc gca tac 5107
Ala His Phe Leu S	Ser Gln Thr Lys	Gln Ser Gly Glu	Asn Phe Ala Tyr
1575	1580	1585	
tta gta gcc tat c	cag god ada gtg	tgc gcc agg gcc	aaa geg eee eee 5155
Leu Val Ala Tyr G	In Ala Thr Val	Cys Ala Arg Ala	Lys Ala Pro Pro
1590	1595	1600	1605
ccg tcc tgg gac g	ite atg tgg aag	tgc ttg act cga	ctc aag ccc acg 5203

Pro Ser Trp Asp Val Met Trp Lys Cys Leu Thr Arg Leu Lys Pro Thr

161	0	1615	1620	
ctt gtg ggc cct				
Leu Val Gly Pro		seu Tyr Arg Le 530	u Giy ser val	Thr Ash
1025	16	,,,0	1033	
gag gtc acc ctt	aca cac ccc g	gtg aca aaa ta	c atc gcc aca	tgc atg 5299
Glu Val Thr Leu	Thr His Pro V	Val Thr Lys Ty	r Ile Ala Thr	Cys Met
1640	1645	3	650	
caa get gae ete				
Gln Ala Asp Leu				Gly Gly
1655	1660	1665		
gtc tta gca gcc	gte gee geg t	tat tgc tta gc	g acc ggg tgt	gtt tcc 5395
Val Leu Ala Ala				
1670	1675	1680	1685	
1670	1675	1680	1685	
1670 atc att ggc egt				
	tta cac atc a	aac cag ega go	et gto gto got	ccg gac 5443
atc att ggc cgt	tta cac atc a	aac cag ega go	et gto gto got	ccg gac 5443
atc att ggc cgt Ile Ile Gly Arg 169	tta cac atc a Leu His Ile A	aac cag cga gc Asn Gln Arg Al 1695	et gte gte get a Val Val Ala 1700	ccg gac 5443 Pro Asp
atc att ggc cgt Tle Tle Gly Arg 169 aag gag gtc ctc	tta cac atc a Leu His Ile A 0 tat gag gct t	aac cag cga go Asn Gln Arg Al 1695 ttt gat gag at	et gtc gtc gct a Val Val Ala 1700	ccg gac 5443 Pro Asp gcc tcc 5491
atc att ggc cgt Ile Ile Gly Arg 169	tta cac atc a Leu His Ile A  0  tat gag gct t Tyr Glu Ala B	aac cag cga go Asn Gln Arg Al 1695 ttt gat gag at	et gtc gtc gct a Val Val Ala 1700	ccg gac 5443 Pro Asp gcc tcc 5491
atc att ggc cgt Tle Tle Gly Arg 169 aag gag gtc ctc Lys Glu Val Leu	tta cac atc a Leu His Ile A  0  tat gag gct t Tyr Glu Ala B	aac cag cga go Asn Gln Arg Al 1695 ttt gat gag at Phe Asp Glu Me	et gtc gtc gct a Val Val Ala 1700 g gag gaa tgt et Glu Glu Cys	ccg gac 5443 Pro Asp gcc tcc 5491
atc att ggc cgt Tle Tle Gly Arg 169 aag gag gtc ctc Lys Glu Val Leu	tta cac atc a Leu His Ile A     tat gag gct t  Tyr Glu Ala E	aac cag cga gc Asn Gln Arg Al 1695 ttt gat gag at Phe Asp Glu Me	et gtc gtc gct a Val Val Ala 1700 g gag gaa tgt et Glu Glu Cys 1715	ccg gac 5443 Pro Asp gcc tcc 5491 Ala Ser
atc att ggc cgt  Ile Ile Gly Arg  169  aag gag gtc ctc Lys Glu Val Leu  1705	tta cac atc a Leu His Ile A  0  tat gag gct t  Tyr Glu Ala B	aac cag cga gc Asn Gln Arg Al 1695 ttt gat gag at Phe Asp Glu Me	et gtc gtc gct a Val Val Ala 1700 g gag gaa tgt et Glu Glu Cys 1715 a gcc gag atg	ccg gac 5443 Pro Asp gcc tcc 5491 Ala Ser ctg aag 5539
atc att ggc cgt Ile Ile Gly Arg 169 aag gag gtc ctc Lys Glu Val Leu 1705	tta cac atc a Leu His Ile A  0  tat gag gct t  Tyr Glu Ala B	aac cag cga gc Asn Gln Arg Al 1695  ttt gat gag at Phe Asp Glu Me 710  ggg cag cgg at	et gtc gtc gct a Val Val Ala 1700 g gag gaa tgt et Glu Glu Cys 1715 a gcc gag atg	ccg gac 5443 Pro Asp gcc tcc 5491 Ala Ser ctg aag 5539
atc att ggc cgt Ile Ile Gly Arg 169 aag gag gtc ctc Lys Glu Val Leu 1705 aga gcg gct ctc Arg Ala Ala Leu 1720	tta cac atc a Leu His Ile A  0  tat gag gct t  Tyr Glu Ala E  17  ctt gaa gag g  Leu Glu Glu G	aac cag cga gc Asn Gln Arg Al 1695  ttt gat gag at Phe Asp Glu Me 710  ggg cag cgg at	et gtc gtc gct a Val Val Ala 1700 g gag gaa tgt et Glu Glu Cys 1715 a gcc gag atg e Ala Glu Met	ccg gac 5443 Pro Asp  gcc tcc 5491 Ala Ser  ctg aag 5539 Leu Lys
atc att ggc cgt Ile Ile Gly Arg 169 aag gag gtc ctc Lys Glu Val Leu 1705 aga gcg gct ctc Arg Ala Ala Leu	tta cac atc a Leu His Ile A  0  tat gag gct t  Tyr Glu Ala E  17  ctt gaa gag g  Leu Glu Glu Glu G  1725	aac cag cga gc Asn Gln Arg Al 1695  ttt gat gag at Phe Asp Glu Me 710  ggg cag cgg at Gly Gln Arg Il	et gtc gtc gct a Val Val Ala 1700  g gag gaa tgt at Glu Glu Cys 1715  a gcc gag atg e Ala Glu Met 1730	ccg gac 5443 Pro Asp  gcc tcc 5491 Ala Ser  ctg aag 5539 Leu Lys  cag gac 5587

1735	174	0	1745	·
Ile Gln Pro	Ala Val Gln		rp Pro Lys Met	gag caa ttc tgg 5635 Glu Gln Phe Trp
1750	1755		1760	1765
•				tac ctc gca gga 5683
Ala Lys His	Met Trp Asn	Phe Ile Se	er Gly Ile Gln	Tyr Leu Ala Gly
	1770	177	5	1780
ctg tca aca	ctg cca ggg	aac cct go	et gtg get tee	atg atg gca ttc 5731
Leu Ser Thr	Leu Pro Gly	Asn Pro Al	la Val Ala Ser	Met Met Ala Phe
178	35	1790	179	5
age gee gee	ctc acc agt	ccg ttg to	ca act age acc	acc atc ctt ctt 5779
Ser Ala Ala	Leu Thr Ser	Pro Leu Se	er Thr Ser Thr	Thr Ile Leu Leu
1800		1805	1810	
aac att ctg	ggg ggc tgg	ctg gcg to	cc caa att gcg	cca ccc gcg ggg 5827
Asn Ile Leu	Gly Gly Trp	Leu Ala Se	er Gln Ile Ala	Pro Pro Ala Gly
1815	182	0	1825	·
gee act gge	ttt gtt gtc	agt ggc ct	tg gtg gga gct	gct gtt ggc agc 5875
Ala Thr Gly	Phe Val Val	Ser Gly Le	eu Val Gly Ala	Ala Val Gly Ser
1830	1835		1840	1845
ata ggc ttg	ggt aaa gtg	ctg gtg ga	ac atc ctg gca	ggg tat ggt gcg 5923
Ile Gly Leu	Gly Lys Val	Leu Val As	sp Ile Leu Ala	Gly Tyr Gly Ala
	1850	185	5	1860
ggc att tcg	ggg gcc ctc	gte geg tt	tt aag atc atg	tct ggc gag aag 5971

Gly Ile Ser Gly Ala Leu Val Ala Phe Lys Ile Met Ser Gly Glu Lys

		18	65				1870				10/	5				
cec	tcc	atg	gag	gat	gtc	atc	aac	ttg	ctg	cct	<b>33</b> 3	att	ctg	tct	cca	6019
Pro	Ser	Меt	G1 u	qaA	۷al	Ile	Asn	Leu	Leu	Pro	Gly	Ile	Leu	Ser	Pro	
	1.8	380				1885				18:	90					
ggt	gct	ctg	gtg	gtg	gga	gtc	atc	tgc	gcg	gcc	att	ctg	cgc	cgc	cat	6067
Gly	Ala	Leu	Val	Val	Gly	Val	Ile	Сув	Ala	Ala	Ile	Leu	Arg	Arg	His	
1	895				190	0			19	05						
gtg	gga	ccg	aaa	gaa	ggc	gcg	gtc	caa	tgg	atg	aac	agg	ctt	atc	gcc	6115
Val	Gly	Pro	Gly	Glu	Gly	Ala	Val	Gln	Trp	Met	Asn	Arg	Leu	Ile	Ala	
1910	)			191	.5			1	920				1925	5		
ttc	gct	tcc	aga	gga	aac	cac	gtc	gcc	cct	act	cac	tac	gtg	acg	gag	6163
Phe	Ala	Ser	Arg	Gly	Asn	His	Val	Ala	Pro	Thr	His	Tyr	Val	Thr	Glu	
			193	0			1	935				1940	)			
tcg	gat	gcg	tcg	cag	cgt	gtc	acc	caa	ctg	ctt	ggc	tct	ctc	act	ata	6211
Ser	Asp	Ala	Ser	Gln	Arg	Val	Thr	Gln	Leu	Leu	Gly	Ser	Leu	Thr	Ile	
		19	45			1	950				195	5				
act	agt	cta	ctc	agg	aga	ctt	cac	aac	tgg	atc	act	gag	gat	tgc	ccc	6259
Thr	Ser	Leu	Leu	Arg	Arg	Leu	His	Asn	Trp	Ile	Thr	Glu	Asp	Cys	Pro	
	19	60				1965				19	70					
atc	сса	tgc	gcc	ggc	tcg	tgg	ctc	cgc	gat	gtg	tgg	gac	tgg	gtc	tgt	6307
Tle	Pro	Cys	Ala	Gly	Ser	Trp	Leu	Arg	Asp	Val	Trp	Asp	Trp	Val	Cys	
1	975				198	)			19	85						
acc	atc	cta	aca	gac	ttt	aag	aac	tgg	ctg	acc	tac	aag	ctg	tte	cca	6355

Thr Ile Leu Thr Asp Phe Lys Asn Trp Leu Thr Ser Lys Leu Phe Pro

1990	)			199	5			2	000				200	5		
aag	atg	cct	ggc	ctc	ccc	ttt	atc	tct	tgc	caa	aag	aaa	tac	aag	ggc	6403
Lys	Met	Pro	Gly	Leu	Pro	Phe	Ile	Ser	Сув	Gln	ГЛЯ	Gly	Tyr	Lys	Gly	
			201	.0			2	015				2020	)			
gtg	tgg	gcc	ggc	act	ggc	atc	atg	acc	aca	cga	tgc	aaa	tgc	ggc	gcc	6451
Val	Trp	Ala	Gly	Thr	Gly	Ile	Met	Thr	Thr	Arg	Сув	Pro	Сув	Gly	Ala	
		20	25			2	2030				203	5				
aac	atc	tct	ggc	aac	gtc	cgc	ttg	ggc	tet	atg	aga	atc	aca	gga	ccc	6499
Asn	Ile	Ser	Gly	Asn	Val	Arg	Leu	Gly	Ser	Met	Arg	Ile	Thr	Gly	Pro	
	2 (	040				2045				20	50					
aaa	acc	ŧgc	atg	aac	acc	tgg	cag	aaa	acc	ttt	act	atc	aat	tgt	tat	6547
Lys	Thr	Сув	Met	Asn	Thr	Trp	Gln	Gly	Thr	Phe	Pro	Ile	Asn	Сув	Tyr	
2	055				206	0			20	65						
2	055				206	0			20	65						
		ggc	cag	tga			aaa	ccc			aac	ttc	aag	acc	gcc	6595
aca	gaa		cag Gln		ttg	ccg			gcg	tta						6595
aca	gaa Glu				ttg Leu	ccg		Pro	gcg	tta				Thr		6595
aca Thr	gaa Glu			Сув	ttg Leu	ccg		Pro	gcg Ala	tta			Lys	Thr		6595
aca Thr 2070	gaa Glu	Gly		Cys 207	ttg Leu	ccg Pro	Lys	Pro 2	gcg Ala 080	tta Leu	Asn	Phe	Lys 208!	Thr	Ala	6595 6643
aca Thr 2070	gaa Glu tgg	Gly aga	Gln	Cys 207 gcg	ttg Leu 5	ccg Pro	Lys gag	Pro 2 tac	gcg Ala 080 gcg	tta Leu gaa	Asn gtg	Phe acg	Lys 208!	Thr	Ala gga	
aca Thr 2070	gaa Glu tgg	Gly aga	Gln gtg	Cys 207 gcg Ala	ttg Leu 5	ccg Pro	Lys gag Glu	Pro 2 tac	gcg Ala 080 gcg	tta Leu gaa	Asn gtg	Phe acg	Lys 2089 cag	Thr	Ala gga	
aca Thr 2070	gaa Glu tgg	Gly aga	Gln gtg Val	Cys 207 gcg Ala	ttg Leu 5	ccg Pro	Lys gag Glu	Pro 2 tac Tyr	gcg Ala 080 gcg	tta Leu gaa	Asn gtg	Phe acg Thr	Lys 2089 cag	Thr	Ala gga	
aca Thr 2070 atc	gaa Glu tgg Trp	Gly aga Arg	gtg Val 209	Cys 207 gcg Ala	ttg Leu 5 gcc Ala	ccg Pro tca Ser	Lys gag Glu 2	Pro 2 tac Tyr 095	gcg Ala 080 gcg Ala	tta Leu gaa Glu	Asn gtg Val	Phe acg Thr 2100	Lys 208! cag Gln	Thr cac His	Ala gga	6643
aca Thr 2070 atc Ile	gaa Glu tgg Trp	Gly aga Arg	gtg Val 209	Cys 207 gcg Ala 0	ttg Leu 5 gcc Ala	ccg Pro tca ser	Lys gag Glu 2 ctg	Pro 2 tac Tyr 095	gcg Ala 080 gcg Ala	tta Leu gaa Glu gac	Asn gtg Val	Phe acg Thr 2100	Lys 2089 cag Gln	Thr cac His	Ala gga Gly	6643
aca Thr 2070 atc Ile	gaa Glu tgg Trp	Gly aga Arg	gtg Val 209 tat	Cys 207 gcg Ala 0	ttg Leu 5 gcc Ala	ccg Pro tca Ser ggg Gly	Lys gag Glu 2 ctg	Pro 2 tac Tyr 095	gcg Ala 080 gcg Ala	tta Leu gaa Glu gac	Asn gtg Val	Phe acg Thr 2100 tta Leu	Lys 2089 cag Gln	Thr cac His	Ala gga Gly	6643
aca Thr 2070 atc Ile	gaa Glu tgg Trp	Gly aga Arg gcc Ala	gtg Val 209 tat	Cys 207 gcg Ala 0	ttg Leu 5 gcc Ala	ccg Pro tca Ser ggg Gly	gag Glu 2 ctg Leu	Pro 2 tac Tyr 095	gcg Ala 080 gcg Ala	tta Leu gaa Glu gac	gtg Val aac Asn	Phe acg Thr 2100 tta Leu	Lys 2089 cag Gln	Thr cac His	Ala gga Gly	6643
aca Thr 2070 atc Ile tca Ser	gaa Glu tgg Trp tat	aga Arg gcc Ala 21	gtg Val 209 tat Tyr	Cys 207 gcg Ala 0 ata Ile	ttg Leu 5 gcc Ala aca Thr	ccg Pro tca Ser ggg Gly	gag Glu 2 ctg Leu	tac Tyr 095 acc	gcg Ala 080 gcg Ala act	tta Leu gaa Glu gac Asp	gtg Val aac Asn 211	Phe acg Thr 2100 tta Leu 5	Lys 208! cag Gln aaa	Thr cac His gtc Val	Ala gga Gly	6643 6691

2120	2125	2130	
		a aag ccg ttt ttc o Lys Pro Phe Phe	cgg gat gag gtc 6787 Arg Asp Glu Val
2135	2140	2145	
teg tte age gtt	ggg etc aat te	a ttt gtc gtc ggg	tot cag ott coc 6835
Ser Phe Ser Val	Gly Leu Asn Se	r Phe Val Val Gly	Ser Gln Leu Pro
2150	2155	2160	2165
tgt gac cct gag	ccc gac act ga	g gta gtg atg tec	atg cta aca gac 6883
Cys Asp Pro Glu	Pro Asp Thr Gl	u Val Val Met Ser	Met Leu Thr Asp
21	70	2175	2180
cca tcc cat atc	acg gcg gag gc	t gca gcg cgg cgt	tta gcg cgg ggg 6931
Pro Ser His Ile	Thr Ala Glu Ala	a Ala Ala Arg Arg	Leu Ala Arg Gly
2185	219	0 21	95
tca ece eca tet	gag gca age te	c tca gcg agc cag	ctg tcg gcg cca 6979
Ser Pro Pro Ser	Glu Ala Ser Se	r Ser Ala Ser Gln	Leu Ser Ala Pro
2200	2205	2210	
tog otg oga goo	acc tgc acc ac	c cac ggt agg acc	tat gat gtg gac 7027
Ser Leu Arg Ala	Thr Cys Thr Th	r His Gly Arg Thr	Tyr Asp Val Asp
2215	2220	2225	
atg gtg gat gcc	aac ctg ttc at	g ggg ggc ggc gtg	att cgg ata gag 7075
Met Val Asp Ala	Asn Leu Phe Me	t Gly Gly Gly Val	Ile Arg Ile Glu
2230	2235	2240	2245
tet gag tee aaa	gtg gtc gtt ctg	g gac tcc ctc gac	tca atg acc gag 7123

Ser Glu Ser Lys Val Val Leu Asp Ser Leu Asp Ser Met Thr Glu

			225	0			2	255				2260	)			
gaa	gag	ggc	gac	ctt	gag	cct	tca	gta	cca	tcg	gag	tat	atg	ctc	ccc	7171
Glu	Glu	Gly	Asp	Leu	Glu	Pro	Ser	Va1	Pro	ser	Glu	Tyr	Met	Leu	Pro	
		22	65			2	2270				227	5				
agg	aag	agg	ttc	cca	ccg	gcc	tta	ccg	gct	tgg	gcg	cgg	cct	gat	tac	7219
Arg	Lys	Arg	Phe	Pro	Pro	Ala	Leu	Pro	Ala	Trp	Ala	Arg	Pro	qeA	Tyr	
	22	280				2285				22	90					
aac	cca	ccg	ctt	gtg	gaa	teg	tgg	aag	agg	cca	gat	tac	caa	cca	ccc	7267
Asn	Pro	Pro	Leu	Val	Glu	Ser	Trp	Lys	Arg	Pro	Asp	Tyr	Gln	Pro	Pro	
2	295				230	0			23	05						
act	gtt	gcg	ggc	tgt	gct	ctc	ccc	ccc	ccc	aaa	aag	acc	ccg	acg	cct	7315
Thr	Val	Ala	Gly	Сув	Ala	Leu	Pro	Pro	Pro	Lys	Lys	Thr	Pro	Thr	Pro	
2310	)			231	. 5			2	320				2325	5		
cct	cca	agg	aga	cgc	cgg	aca	gtg	ggt	ctg	agc	gag	agc	acc	ata	gga	7363
Pro	Pro	Arg	Arg	Arg	Arg	Thr	Val	Gly	Leu	Ser	Glu	Ser	Thr	Ile	Gly	
			233	0			2	335				2340	0			
gat	gcc	ctc	caa	cag	ctg	gcc	atc	aag	tcc	ttt	ggc	cag	ccc	ccc	cca	7411
Asp	Ala	Leu	Gln	Gln	Leu	Ala	Ile	Lys	Ser	Phe	Gly	Gln	Pro	Pro	Pro	
		23	45			2	2350				235	5				
agc	ggc	gat	tca	ggc	ctt	tcc	acg	<b>3</b> 33	gcg	gac	gcc	gcc	gac	tcc	ggc	7459
ser	Gly	Asp	Ser	Gly	Leu	Ser	Thr	Gly	Ala	Asp	Ala	Ala	Asp	Ser	Gly	
	23	860				2365				23	70					

gat egg aca ece eet gae gag ttg get ett teg gag aca ggt tet ace Asp Arg Thr Pro Pro Asp Glu Leu Ala Leu Ser Glu Thr Gly Ser Thr

2375	238	0	2385		
tee tee atg	ccc ccc ctc	gag ggg gag	cct ggg gac	cca gac ctg	gag 7555
Ser Ser Met	Pro Pro Leu	Glu Gly Glu	Pro Gly Asp	Pro Asp Leu	Glu
2390	2395	2	400	2405	
cct gag cag	gta gag ctt	caa cct cct	ccc cag ggg	ggg gag gca	gct 7603
Pro Glu Gln	Val Glu Leu	Gln Pro Pro	Pro Gln Gly	Gly Glu Ala	Ala
	2410	2415		2420	
acc ggc tag	gac teg ggg	tee tgg tet	act tgc tcc	gag gag gat	gac 7651
Pro Gly Ser	Asp Ser Gly	Ser Trp Ser	Thr Cys Ser	Glu Glu Asp	Asp
243	25	2430	243	5	
tee gte gtg	tgc tgc tcc	atg tca tat	tec tgg acc	ggg gct cta	ata 7699
Ser Val Val	Cys Cys Ser	Met Ser Tyr	Ser Trp Thr	Gly Ala Leu	Ile
2440		2445	2450		
act cct tgt	agc ccc gaa	gag gaa aag	ttg cca att	aac tcc ttg	agc 7747
Thr Pro Cys	Ser Pro Glu	Glu Glu Lys	Leu Pro Ile	Asn Ser Leu	Ser
2455	246	0	2465		
aac tog otg	ttg cga tac	cat aac aag	gta tac tgt	act aca tca	aag 7795
Asn Ser Leu	Leu Arg Tyr	His Asn Lys	Val Tyr Cys	Thr Thr Ser	Lys
2470	2475	2	480	2485	
agt gcc tca	cta agg gct	aaa aag gta	act ttt gat	agg atg caa	gtg 7843
Ser Ala Ser	Leu Arg Ala	Lys Lys Val	Thr Phe Asp	Arg Met Gln	Val
	2490	2495		2500	

ctc gac gcc tat tat gat tca gtc tta aag gac atc aag cta gcg gcc Leu Asp Ala Tyr Tyr Asp Ser Val Leu Lys Asp Ile Lys Leu Ala Ala

25	05	:	2510			251	5				
tcc aag gtc											7939
Ser Lys Val	Ser Ala	Arg Leu	Leu Th	hr Leu	Glu	Glu	Ala	Cys	Gln	Leu	
2520		2525			253	3 0					
ace cca ecc	cac tct	gca aga	tcc aa	ag tat	ggg	ttt	999	gct	aag	gag	7987
Thr Pro Pro	His Ser	Ala Arg	Ser Ly	ys Tyr	Gly	Phe	Gly	Ala	Lys	Glu	
2535		2540		25	45						
gtc cgc agc	tta tee	ggg agg	acc at	tc aac	cac	atc	aaq	tee	gtg	tgg	8035
Val Arg Ser											
			nia ve	2560	*****		_,,	2569			
2550	255	) \$		4560				250.	,		
aag gac ctc	ttg gaa	gac tca	caa a	ca cca	att	cct	aca	acc	atc	atg	8083
Lys Asp Leu	Leu Glu	Asp Ser	Gln Tl	hr Pro	Ile	Pro	Thr	Thr	Ile	Met	
	2570		257	5			2580	)			
gec aaa aat	gag gtg	ttc tgc	gtg ga	ac ccc	gcc	aag	ggg	ggt	aaa	aaa	8131
Ala Lys Asn	Glu Val	Phe Cys	Val As	sp Pro	Ala	Lys	Gly	Gly	Lys	Lys	
25	85	:	2590			259	5				
cca gct cgc	off ato	att tac	cct a	ac ctc	aac	ate	ааа	atic	tac	gag	8179
Pro Ala Arg											
_	neu iie	_		эр пец			vrā	vai	Cys	OIU	
2600		2605	•		260	LO					
aag atg gcc	ctt tat	gat gtc	aca ca	aa aag	ctt	cct	cag	gcg	gtg	atg	8227
Lys Met Ala	Leu Tyr	Asp Val	Thr G	ln Lys	Leu	Pro	Gln	Ala	Val	Met	
2615		2620		26	25						
ggg gct tct	tat ggc	ttc cag	tac to	cc ccc	gct	cag	cgg	gtg	gag	ttt	8275

Gly Ala Ser Tyr Gly Phe Gln Tyr Ser Pro Ala Gln Arg Val Glu Phe

2630	)			263	5			2	640				2645	j .		
ata	++~	224	aan.	+44	~~~	~22	220	202	asc	aat	ato	aat		tca	tat	8323
			Ala													
ьец	Бец	цув	265		nia	Gru		655	nop	***	1700	2660		501	-1-	
			200				-	033				200				
gat	acc	ega	tgc	ttt	gac	tca	acc	gtc	act	gag	aga	gac	atc	agg	act	8371
Asp	Thr	Arg	Cys	Phe	Asp	ser	Thr	Val	Thr	Glu	Arg	Asp	Ile	Arg	Thr	
		26	65			2	2670				267	5				
gag	gag	tee	ata	tac	cag	gcc	tgc	tcc	tta	ccc	gag	gag	gcc	cga	act	8419
Glu	Glu	Ser	Ile	Tyr	Gln	Ala	Cys	Ser	Leu	Pro	Glu	Glu	Ala	Arg	Thr	
	26	580				2685				26	90					
gcc	ata	cac	tcg	ctg	act	gag	aga	ctc	tat	gtg	gga	999	ccc	atg	ttc	8467
Ala	Ile	His	Ser	Leu	Thr	Glu	Arg	Leu	Tyr	Val	Gly	Gly	Pro	Met	Phe	
2	695				270	0			27	05						
aac	agc	aag	ggc	cag	tcc	tgc	999	tac	agg	cgt	tgc	cgc	gcc	agc	<b>3</b> 33	8515
Asn	Ser	Lys	Gly	Gln	Ser	Сув	Gly	Tyr	Arg	Arg	Сув	Arg	Ala	Ser	Gly	
271	)			271	. 5			2	720				272	5		
gtg	ctt	acc	act	agt	atg	aaa	aac	acc	atc	aca	tgc	tat	gta	aaa	gcc	8563
Val	Leu	Thr	Thr	Ser	Met	Gly	Asn	Thr	Ile	Thr	Сув	Tyr	Val	Lys	Ala	
			273	0			2	735				274	0			
cta	gcg	gct	tgc	aag	g¢t	gcg	ggg	ata	att	gcg	ecc	acg	atg	ctg	gta	8611
Leu	Ala	Ala	Cys	Lys	Ala	Ala	Gly	Ile	Ile	Ala	Pro	Thr	Met	Leu	Val	
		27	45			:	2750				275	55				
tgc	ggc	gac	gac	ttg	gtc	gtc	atc	tca	gaa	agc	cag	ggg	act	gag	gag	8659
Cve	Gly	Agn	apa	Leu	Val	Val	Tle	Ser	Glu	Ser	Gln	Glv	Thr	Glu	Glu	

gac gag cgg aac	ctg aga gcc	ttc acg gag	get atg acc	agg tat tct 8707
Asp Glu Arg Asn	Leu Arg Ala	Phe Thr Glu	Ala Met Thr	Arg Tyr Ser
2775	2780	27	85	
gee eet eet ggt	gac ccc ccc	aga ccg gaa	tat gac ctg	gag cta ata 8755
Ala Pro Pro Gly	Asp Pro Pro	Arg Pro Glu	Tyr Asp Leu	Glu Leu Ile
2790	2795	2800		2805
aca tot tgt too	tca aac gtg	tet gtg gea	ctt ggc cca	cag ggc cgc 8803
Thr Ser Cys Ser	Ser Asn Val	Ser Val Ala	Leu Gly Pro	Gln Gly Arg
28	10	2815	282	0
cgc aga tac tac	ctg acc aga	gac ccc acc	act tca att	gcc cgg gct 8851
Arg Arg Tyr Tyr	Leu Thr Arg	Asp Pro Thr	Thr Ser Ile	Ala Arg Ala
2825	2	2830	2835	
gcc tgg gaa aca	gtt aga cac	tee eet ate		ctq qqa aac 8899
		cee eee gee	aat tea tgg	229 354 444
Ala Trp Glu Thr	Val Arg His			-
Ala Trp Glu Thr 2840	Val Arg His	Ser Pro Val		-
-	-	Ser Pro Val	Asn Ser Trp	-
-	2845	Ser Pro Val	Asn Ser Trp 2850	Leu Gly Asn
2840	2845 get eca ace	Ser Pro Val	Asn Ser Trp 2850 cgc atg gtc	Leu Gly Asn ctg atg aca 8947
2840 atc atc cag tac	2845 get eca ace	Ser Pro Val  ata tgg gtt  Ile Trp Val	Asn Ser Trp 2850 cgc atg gtc	Leu Gly Asn ctg atg aca 8947
2840 atc atc cag tac Ile Ile Gln Tyr	2845 gct cca acc Ala Pro Thr	Ser Pro Val  ata tgg gtt  Ile Trp Val	Asn Ser Trp 2850 cgc atg gtc Arg Met Val	Leu Gly Asn ctg atg aca 8947
2840 atc atc cag tac Ile Ile Gln Tyr	get cca acc Ala Pro Thr 2860	Ser Pro Val ata tgg gtt Ile Trp Val 28	Asn Ser Trp 2850 cgc atg gtc Arg Met Val	Leu Gly Asn  ctg atg aca 8947  Leu Met Thr
2840  atc atc cag tac  Ile Ile Gln Tyr  2855	gct cca acc Ala Pro Thr 2860 att ctc atg	ser Pro Val  ata tgg gtt  Ile Trp Val  28  gcc cag gac	Asn Ser Trp 2850  cgc atg gtc Arg Met Val 65  acc cta gac	Leu Gly Asn  ctg atg aca 8947  Leu Met Thr  cag aac ctt 8995
2840  atc atc cag tac  Ile Ile Gln Tyr  2855  cac ttc ttc tcc	gct cca acc Ala Pro Thr 2860 att ctc atg	ser Pro Val  ata tgg gtt  Ile Trp Val  28  gcc cag gac	Asn Ser Trp 2850  cgc atg gtc Arg Met Val 65  acc cta gac	Leu Gly Asn  ctg atg aca 8947  Leu Met Thr  cag aac ctt 8995

2760

2765

2770

aac ttt gaa atg tac gga tcg gtg tac tcc gtg agt cct ctg gac ctc 9043

Asn Phe Glu Met Tyr Gly Ser Val Tyr Ser Val Ser Pro Leu Asp Leu

				289	0			2	895				2900	)			
																cac	9091
P	ΓŲ	Ala	29		Glu	Arg		2910	GIY	neu	Asp	291		per	пец	пта	
a	ca	tac	act	aaa	cac	gaa	ctg	acg	cgg	gtg	gct	tca	gcc	ctc	aga	aaa	9139
т	hr	Tyr	Thr	Pro	His	Glu	Leu	Thr	Arg	Val	Ala	Ser	Ala	Leu	Arg	ГАв	
		25	920				2925				29	3 0					
_								~~~		224	200		~~~	aa+	~~~	att	9197
					Pro												9187
ы		935	nau	110	110	294		712 G	111		45	5		***************************************			
a	99	gcg	tcc	ctc	atc	tec	cgt	aaa	aaa	agg	gcg	gcc	gtt	tgc	ggt	cgg	9235
A	rg	Ala	ser	Leu	Ile	Ser	Arg	Gly	Gly	Arg	Ala	Ala	Val	Сув	Gly	Arg	
2	95(	)			295	55			2	960				296	õ		
																	0202
t	ac	ctc			tgg	gcg			acc	aag				act	cct	t t g Leu	9283
t	ac	ctc			tgg Trp	gcg		Lys	acc	aag				act Thr	cct		9283
t	ac	ctc		Asn	tgg Trp	gcg		Lys	acc Thr	aag			Leu	act Thr	cct		9283
t	ac yr	ctc Leu	Phe	Asn 297	tgg Trp	gcg Ala	Val	Lys 2	acc Thr 975	aag Lys	Leu	Lys	Leu 2980	act Thr	cct Pro		9283 9331
t T	ac yr cg	ctc Leu gag	Phe gca	Asn 297 ege	tgg Trp	geg Ala	Val gat	Lys 2 ttg	acc Thr 975 tcc	aag Lys agt	Leu tgg	Lys	Leu 2980 acc	act Thr	cct Pro	Leu gcc	
t T	ac yr cg	ctc Leu gag	Phe gca Ala	Asn 297 ege	tgg Trp 0	geg Ala	Val gat Asp	Lys 2 ttg	acc Thr 975 tcc	aag Lys agt	Leu tgg	Lys	Leu 2980 acc	act Thr	cct Pro	Leu gcc	
t T	ac yr cg ro	ctc Leu gag Glu	Phe gca Ala 29	Asn 297 ege Arg 85	tgg Trp 0 ctc Leu	gcg Ala ctg Leu	Val gat Asp	Lys 2 ttg Leu 2990	acc Thr 975 tcc Ser	aag Lys agt ser	Leu tgg Trp	Lys ttt Phe 299	Leu 2980 acc Thr	act Thr ) gtc Val	ect Pro ggc Gly	Leu gcc Ala	9331
t T c	ac yr cg ro	ctc Leu gag Glu	gca Ala 29	Asn 297 egc Arg 85	tgg Trp 70 ctc Leu	gcg Ala ctg Leu	Val gat Asp	ttg teu 1990 agc	acc Thr 975 tcc Ser	aag Lys agt ser	tgg Trp	ttt Phe 299	Leu 2980 acc Thr	act Thr  gtc Val	cct Pro ggc Gly	Leu gcc Ala	
t T c	ac yr cg ro	ctc Leu gag Glu ggg Gly	gca Ala 29	Asn 297 egc Arg 85	tgg Trp 0 ctc Leu	gcg Ala ctg Leu	Val gat Asp	ttg teu 1990 agc ser	acc Thr 975 tcc Ser	aag Lys agt ser	tgg Trp	ttt Phe 299 gcc	Leu 2980 acc Thr	act Thr  gtc Val	cct Pro ggc Gly	Leu gcc Ala	9331
t T c	ac yr cg ro	ctc Leu gag Glu ggg Gly	gca Ala 29 ggc	Asn 297 egc Arg 85	tgg Trp 70 ctc Leu	gcg Ala ctg Leu	yat Asp cac	ttg teu 1990 agc ser	acc Thr 975 tcc Ser	aag Lys agt ser	tgg Trp cgt	ttt Phe 299 gcc	Leu 2980 acc Thr	act Thr  gtc Val	cct Pro ggc Gly	Leu gcc Ala	9331
t T C P	ac yr cg ro	ctc Leu gag Glu ggg Gly	gca Ala 29 ggc Gly	Asn 297 egc Arg 85 gac Asp	tgg Trp 70 ctc Leu	gcg Ala ctg Leu tat	yat Asp cac His	ttg teu Leu 2990 agc ser	acc Thr 975 tcc Ser gtg Val	aag Lys agt ser tcg ser	tgg Trp cgt Arg	ttt Phe 299 gcc Ala	Leu 2980 acc Thr 5	act Thr gtc Val	ect Pro ggc Gly egc Arg	gcc Ala cta	9331

3015 3020 3025

ctc ccc gct cga tag agcggcacac attagctaca ctccatagct aactgttcct 9482 Leu Pro Ala Arg

3030

theirtett terrettt terrettet terrettet terrettet terrettet terrettet terrectet 9542

tettecette teatettatt etaetttett tettggtgge tecatettag ecetggteac 9602

ggctagctgt gaaaggtccg tgagccgcat gactgcagag agtgccgtaa ctggtctctc 9662

tgcagatcat gt 9674

<210> 6

<211> 3033

<212> PRT

<213> Hepatitis C virus

<400> 6

Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn

1 5 10 15

Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly

20 25 30

Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala

35 40 45

Thr Arg Lys Ala Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro

50 55 60

Ile Pro Lys His Arg Arg Ser Thr Gly Lys Ser Trp Gly Lys Pro Gly

65 70 75 80

Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp

			8	5				90				95			
Leu	Leu	Ser	Pro	Arg	Gly	Ser	Arg	Pro	Ser	Trp	Gly	Pro	Asn	Asp	Pro
		1	00				105				11	0			
Arg	His	Arg	Ser	Arg	Asn	Val	G1y	Lys	Val	Ile	Asp	Thr	Leu	Thr	Сув
	1	.15				120				12	5				
Gly	Phe	Ala	Asp	Leu	Leu	Gly	Tyr	Val	Pro	Val	Val	Gly	Ala	Pro	Leu
	130				135	;			1	40					
Ser	Gly	Val	Ala	Ser	Ala	Leu	Ala	His	Gly	Val	Arg	Val	Leu	Glu	Asp
145				15	0			1	155				160		
Gly	Val	Asn	Phe	Ala	Thr	Gly	Asn	Leu	Pro	Gly	Cys	Ser	Phe	Ser	Ile
			16	5				170				175			
Phe	Leu	Leu	Ala	Leu	Leu	ser	Cys	Ile	Thr	Thr	Pro	Val	Ser	Ala	Val
		1	80				185				19	0			
Gln	Val	Lys	Asn	Thr	Ser	Asn	Ala	Tyr	Met	Ala	Thr	Asn	Asp	Сув	Ser
	1	.95				200				20	5				
Asn	Asp	Ser	Ile	Thr	Trp	Gln	Leu	Glu	Ala	Ala	Val	Leu	His	Val	Pro
	210				219	5			2	20					
		Val	Pro	Cys			Met	Gly			Ser	Arg	Сув	Trp	Ile
		Val	Pro	Cys 23	Glu		Met				Ser	Arg	Cys 240	Trp	Ile
Gly 225	Сув			23	Glu 0	Lys			Asn 235	Thr			240		
Gly 225	Сув			23 Asn	Glu 0	Lys	Val	2	Asn 235	Thr			240 Leu		
Gly 225 Pro	Cys Val	Ser	Pro 24	23 Asn 5	Glu 0 Val	Lys Ala	Val	Arg	Asn 235 Gln	Thr	Gly	Ala 255	240 Leu	Thr	Arg
Gly 225 Pro	Cys Val	Ser Arg	Pro 24	23 Asn 5	Glu 0 Val	Lys Ala	Val	Arg 250	Asn 235 Gln	Thr	Gly	Ala 255 Ala	240 Leu	Thr	Arg
Gly 225 Pro	Cys Val Leu	Ser Arg 2	Pro 24 Thr	23 Asn 5 His	Glu 0 Val Ile	Lys Ala Asp	Val Met 265	Arg 250	Asn 235 Gln Val	Thr Pro	Gly Ser 27	Ala 255 Ala 0	240 Leu Thr	Thr Leu	Arg
Gly 225 Pro	Cys Val Leu Ala	Ser Arg 2	Pro 24 Thr	23 Asn 5 His	Glu 0 Val Ile	Lys Ala Asp	Val Met 265	Arg 250 Val	Asn 235 Gln Val	Thr Pro	Gly Ser 27 Val	Ala 255 Ala 0	240 Leu Thr	Thr Leu	Arg
Gly 225 Pro Gly Ser	Cys Val Leu Ala	Ser Arg 2 Leu	Pro 24 Thr 60 Tyr	23 Asn 5 His	Glu 0 Val Ile Gly	Lys Ala Asp Asp 280	Val Met 265 Leu	Arg 250 Val	Asn 235 Gln Val Gly	Thr Pro Leu Gly	Gly Ser 27 Val	Ala 255 Ala 0 Met	240 Leu Thr	Thr Leu Ala	Arg Cys
Gly 225 Pro Gly Ser	Cys Val Leu Ala	Ser Arg 2 Leu	Pro 24 Thr 60 Tyr	23 Asn 5 His	Glu 0 Val Ile Gly	Ala Asp Asp 280 Pro	Val Met 265 Leu	Arg 250 Val Cys	Asn 235 Gln Val Gly	Thr Pro Leu Gly	Gly Ser 27 Val	Ala 255 Ala 0 Met	240 Leu Thr	Thr Leu Ala	Arg Cys
Gly 225 Pro Gly Ser	Cys Val Leu Ala 2 Met	Ser Arg 2 Leu 275	Pro 24 Thr 60 Tyr	23 Asn 5 His Val	Glu  O  Val  Ile  Gly  Ser  295	Ala Asp Asp 280 Pro	Val Met 265 Leu Gln	Arg 250 Val Cys	Asn 235 Gln Val Gly His	Thr Pro Leu Gly 28 Trp	Gly Ser 27 Val S5	Ala 255 Ala 0 Met Val	240 Leu Thr Leu Gln	Thr Leu Ala Glu	Arg Cys Ser
Gly 225 Pro Gly Ser	Cys Val Leu Ala 2 Met	Ser Arg 2 Leu 275	Pro 24 Thr 60 Tyr	23 Asn 5 His Val	Glu  Val  Ile  Gly  Ser  295	Ala Asp Asp 280 Pro	Val Met 265 Leu Gln	Arg 250 Val Cys His	Asn 235 Gln Val Gly His	Thr Pro Leu Gly 28 Trp	Gly Ser 27 Val S5	Ala 255 Ala 0 Met Val	240 Leu Thr Leu Gln	Thr Leu Ala Glu Ala	Arg Cys Ser
Gly 225 Pro Gly Ser Gln Asn 305	Val Leu Ala 2 Met 290 Cys	Ser Arg 2 Leu 75 Phe	Pro 24 Thr 60 Tyr Ile	23 Asn 5 His Val Val Tyr 31	Glu  Val  Ile  Gly  Ser  295  Pro 0	Ala Asp Asp 280 Pro	Val Met 265 Leu Gln	Arg 250 Val Cys His	Asn 235 Gln Val Gly His 3 Thr	Thr Pro Leu Gly 28 Trp 00 Gly	Gly Ser 27 Val S5 Phe	Ala 255 Ala 0 Met Val	240 Leu Thr Leu Gln Met 320	Thr Leu Ala Glu Ala	Arg Cys Ser Cys
Gly 225 Pro Gly Ser Gln Asn 305	Val Leu Ala 2 Met 290 Cys	Ser Arg 2 Leu 75 Phe	Pro 24 Thr 60 Tyr Ile	23 Asn 5 His Val Val Tyr 31 Asn	Glu  Val  Ile  Gly  Ser  295  Pro 0	Ala Asp Asp 280 Pro	Val Met 265 Leu Gln Ala	Arg 250 Val Cys His	Asn 235 Gln Val Gly His 3 Thr	Thr Pro Leu Gly 28 Trp 00 Gly	Gly Ser 27 Val S5 Phe	Ala 255 Ala 0 Met Val	240 Leu Thr Leu Gln Met 320 Leu	Thr Leu Ala Glu Ala	Arg Cys Ser Cys

		3	40				345				3 5	0			
Trp	Gly	Val	Met	Phe	Gly	Leu	Ala	Tyr	Phe	Ser	Met	Gln	Gly	Ala	Trp
	3	55				360				3 6	55				
Ala	Lys	Val	Val	Val	Ile	Leu	Leu	Leu	Ala	Ser	Gly	Val	Asp	Ala	Tyr
	370				375	5			3	80					
Thr	Thr	Thr	Thr	Gly	Ser	Ala	Ala	Gly	Arg	Thr	Thr	Ser	Ser	Leu	Ala
385				3 9	0			3	395				400		
Ser	Ala	Phe	Ser	Pro	Gly	Ala	Arg	Gln	Asn	Ile	Gln	Leu	Ile	Asn	Thr
			4 0	5				410				415			
Asn	Gly	ser	Trp	His	Ile	Asn	Arg	Thr	Ala	Leu	Asn	Сув	Asn	Asp	Ser
		4	20				425				43	0			
Leu	His	Thr	Gly	Phe	Phe	Thr	Ala	Leu	Phe	Tyr	Ile	His	Lys	Phe	Asn
	4	35				440				4 4	5				
Ser	Ser	Gly	Сув	Pro	Glu	Arg	Leu	Ser	Ala	Сув	Arg	Asn	Ile	Glu	Asp
	450				455	5			4	60					
Phe	Arg	Ile	Gly	Trp	Gly	Ala	Leu	Gln	Tyr	Asp	Asp	Asn	Val	Thr	Asn
465				47	0			4	175				480		
	Glu	qaA	Met			Tyr	Сув			Tyr	Pro	Pro		Gln	Суз
	Glu	Asp	Met 48	Arg		Tyr				Туг	Pro	Pro 495	Lys	Gln	Cys
Pro	Glu Val		48	Arg	Pro			Trp 490	His			495	Lys		
Pro		Val	48	Arg	Pro			Trp 490	His			495 Tyr	Lys		
Pro Gly		Val	48 Pro 00	Arg 5 Ala	Pro	Thr	Val 505	Trp 490 Cys	Нів	Pro	Val 51	495 Tyr 0	Lув Сув	Phe	Thr
Pro Gly	Val Ser	Val	48 Pro 00	Arg 5 Ala	Pro	Thr	Val 505	Trp 490 Cys	Нів	Pro	Val 51 Leu	495 Tyr 0	Lув Сув	Phe	Thr
Pro Gly Pro	Val Ser	Val 5 Pro 15	48 Pro 00 Val	Arg 5 Ala Val	Pro Gly Val	Thr Gly 520	Val 505 Thr	Trp 490 Cys Thr	His Gly Asp	Pro Arg	Val 51 Leu 5	495 Tyr 0 Gly	Lys Cys Val	Phe Pro	Thr
Pro Gly Pro Tyr	Val Ser	Val 5 Pro 15	48 Pro 00 Val	Arg 5 Ala Val	Pro Gly Val	Thr Gly 520 Glu	Val 505 Thr	Trp 490 Cys Thr	His Gly Asp Val	Pro Arg	Val 51 Leu 5	495 Tyr 0 Gly	Lys Cys Val	Phe Pro	Thr
Pro Gly Pro Tyr	Val Ser 5	Val 5 Pro 15 Trp	48 Pro 00 Val Gly	Arg 5 Ala Val Glu	Pro Gly Val Asn	Thr Gly 520 Glu	Val 505 Thr	Trp 490 Cys Thr	His Gly Asp Val	Pro Arg 52 Phe 40	Val 51 Leu 5	495 Tyr 0 Gly Leu	Lys Cys Val Asn	Phe Pro Ser	Thr Thr
Pro Gly Pro Tyr	Val Ser 5 Thr	Val 5 Pro 15 Trp	48 Pro 00 Val Gly	Arg 5 Ala Val Glu	Pro Gly Val Asn 539	Thr Gly 520 Glu	Val 505 Thr	Trp 490 Cys Thr Asp	His Gly Asp Val	Pro Arg 52 Phe 40	Val 51 Leu 5	495 Tyr 0 Gly Leu	Lys Cys Val Asn	Phe Pro Ser	Thr Thr
Pro Gly Pro Tyr Arg 545	Val Ser 5 Thr	Val 5 Pro 15 Trp	48 Pro 00 Val Gly Ser	Arg 5 Ala Val Glu Gly 55	Gly Val Asn 535 Ser	Thr Gly 520 Glu Trp	Val 505 Thr Thr	Trp 490 Cys Thr Asp	His Gly Asp Val 5 Cys	Pro Arg 52 Phe 40 Thr	Val 51 Leu 5 Leu Trp	495 Tyr O Gly Leu Met	Lys Cys Val Asn Asn	Phe Pro Ser	Thr Thr Thr
Pro Gly Pro Tyr Arg 545	Val Ser 5 Thr 530 Pro	Val 5 Pro 15 Trp	48 Pro 00 Val Gly Ser	Arg  Ala  Val  Glu  Gly  55  Thr	Gly Val Asn 535 Ser	Thr Gly 520 Glu Trp	Val 505 Thr Thr	Trp 490 Cys Thr Asp	His Gly Asp Val 5 Cys	Pro Arg 52 Phe 40 Thr	Val 51 Leu 5 Leu Trp	495 Tyr O Gly Leu Met	Lys Cys Val Asn Asn 560 Arg	Phe Pro Ser	Thr Thr Thr
Pro Gly Pro Tyr Arg 545 Gly	Val Ser 5 Thr 530 Pro	Val 5 Pro 15 Trp Pro	48 Pro 00 Val Gly Ser Lys	Arg 5 Ala Val Glu Gly 55 Thr	Pro Gly Val Asn 539 Ser 0 Cys	Thr  Gly  520  Glu  Trp	Val 505 Thr Thr Phe	Trp 490 Cys Thr Asp Gly 970	His Gly Asp Val 5 Cys 555	Pro Arg 52 Phe 40 Thr	Val 51 Leu 5 Leu Trp	495 Tyr 0 Gly Leu Met Thr	Lys Cys Val Asn 560 Arg	Pro Ser Ser	Thr Thr Thr
Pro Gly Pro Tyr Arg 545 Gly	Val Ser 5 Thr 530 Pro	Val 5 Pro 15 Trp Pro Thr	48 Pro 00 Val Gly Ser Lys	Arg 5 Ala Val Glu Gly 55 Thr	Pro Gly Val Asn 539 Ser 0 Cys	Thr  Gly  520  Glu  Trp	Val 505 Thr Thr Phe	Trp 490 Cys Thr Asp Gly 970	His Gly Asp Val 5 Cys 555	Pro Arg 52 Phe 40 Thr	Val 51 Leu 5 Leu Trp	495 Tyr 0 Gly Leu Met Thr 575	Lys Cys Val Asn 560 Arg	Pro Ser Ser	Thr Thr Thr

	í	595				600				60	5				
Pro	Lys	Сув	Leu	Val	Asp	Tyr	Pro	Туг	Arg	Leu	Trp	His	Туг	Pro	Сув
	610				61!	ភ			6	20					
Thr	Val	Asn	Tyr	Ser	Thr	Phe	Lys	Ile	Arg	Met	Tyr	Val	Gly	Gly	Val
625				63	0			(	635				640		
Glu	His	Arg	Leu	Met	Ala	Ala	Cys	Asn	Phe	Thr	Arg	Gly	Asp	Arg	Cys
			64	15				650				655			
Asn	Leu	Glu	Asp	Arg	Asp	Arg	Ser	Gln	Gln	Thr	Pro	Leu	Leu	His	Ser
		6	60				665				67	0			
Thr	Thr	Glu	Trp	Ala	Ile	Leu	Pro	Cys	Ser	Phe	Ser	Asp	Leu	Pro	Ala
	6	575				680				68	5				
Leu	Ser	Thr	Gly	Leu	Leu	His	Leu	His	Gln	Asn	Ile	Va1	Asp	Val	Gln
	690				695	5			7	00					
Tyr	Met	Tyr	Gly	Leu	Ser	Pro	Ala	Leu	Thr	Gln	Tyr	Ile	Val	Arg	Trp
705				71	0			•	715				720		
Glu	Trp	Val	Val	Leu	Leu	Phe	Leu	Leu	Leu	Ala	Asp	Ala	Arg	Va1	Сув
			72	:5				730				735			
Ala	Сув	Leu		5 Met	Leu	Ile			Gly	Gln	Ala			Ala	Leu
Ala	Cys				Leu	Ile			Gly	Gln	Ala 75	Glu		Ala	Leu
		7	Trp 40				Leu 745	Leu	_		75	Glu 0	Ala		
	Lys	7	Trp 40	Met			Leu 745	Leu	_		75 Ala	Glu 0	Ala		
Glu	Lys	7. Leu 755	Trp 40 Val	Met	Leu	His 760	Leu 745 Ala	Leu Ala	Ser	Ala 76	75 Ala 5	Glu 0 Ser	Ala Сув	Asn	Gly
Glu Phe	Lys	7. Leu 755	Trp 40 Val	Met Val	Leu	His 760 Phe	Leu 745 Ala	Leu Ala	Ser	Ala 76	75 Ala 5	Glu 0 Ser	Ala Сув	Asn	Gly
Glu Phe	Lys Leu 770	7. Leu 755 Tyr	Trp 40 Val Phe	Met Val	Leu Ile 775	His 760 Phe	Leu 745 Ala Leu	Leu Ala Val	Ser Ala	Ala 76 Ala 80	75 Ala 5 Trp	Glu 0 Ser His	Ala Cys Ile	Asn Lys	Gly Gly
Glu Phe	Lys Leu 770	7 Leu 755 Tyr	Trp 40 Val Phe	Met Val Val	Leu Ile 775 Ala	His 760 Phe	Leu 745 Ala Leu	Leu Ala Val Ser	Ser Ala	Ala 76 Ala 80	75 Ala 5 Trp	Glu 0 Ser His	Ala Cys Ile	Asn Lys	Gly Gly
Glu Phe Arg 785	Lys Leu 770 Val	Leu 755 Tyr Val	Trp 40 Val Phe Pro	Met Val Val	Leu Ile 775 Ala	His 760 Phe	Leu 745 Ala Leu Tyr	Leu Ala Val Ser	Ser Ala 7 Leu	Ala 76 Ala 80 Thr	75 Ala 5 Trp Gly	Glu O Ser His Leu	Ala Cys Ile Trp	Asn Lys Pro	Gly Gly Phe
Glu Phe Arg 785	Lys Leu 770 Val	Leu 755 Tyr Val	Trp 40 Val Phe Pro	Met Val Val Leu 79	Leu Ile 775 Ala	His 760 Phe	Leu 745 Ala Leu Tyr	Leu Ala Val Ser	Ser Ala 7 Leu	Ala 76 Ala 80 Thr	75 Ala 5 Trp Gly	Glu O Ser His Leu	Ala Cys Ile Trp 800 Tyr	Asn Lys Pro	Gly Gly Phe
Glu Phe Arg 785 Cys	Lys Leu 770 Val Leu	7. Leu 755 Tyr Val Leu	Trp 40 Val Phe Pro Leu 80	Met Val Val Leu 79	Leu Ile 775 Ala 0 Ala	His 760 Phe Ala Leu	Leu 745 Ala Leu Tyr	Leu Ala Val Ser Gln 310	Ser Ala 7 Leu 795 Gln	Ala 76 Ala 80 Thr	75 Ala 5 Trp Gly	Glu 0 Ser His Leu Ala 815	Ala Cys Ile Trp 800	Asn Lys Pro Asp	Gly Gly Phe
Glu Phe Arg 785 Cys	Lys Leu 770 Val Leu	Tyr Val Leu	Trp 40 Val Phe Pro Leu 80	Met Val Val Leu 79 Leu 5	Leu Ile 775 Ala 0 Ala	His 760 Phe Ala Leu	Leu 745 Ala Leu Tyr	Leu Ala Val Ser Gln 310	Ser Ala 7 Leu 795 Gln	Ala 76 Ala 80 Thr	75 Ala 5 Trp Gly	Glu  Ser  His  Leu  Ala  815	Ala Cys Ile Trp 800	Asn Lys Pro Asp	Gly Gly Phe
Glu Phe Arg 785 Cys	Lys Leu 770 Val Leu Val	Leu 755 Tyr Val Leu His	Trp 40 Val Phe Pro Leu 80 Gly	Met Val Val Leu 79 Leu 5	Leu Ile 775 Ala O Ala Val	His 760 Phe Ala Leu	Leu 745 Ala Leu Tyr Pro Ala 825	Leu Ala Val Ser Gln 310 Ala	Ser Ala 7 Leu 795 Gln Leu	Ala 76 Ala 80 Thr Ala Leu	756 Ala 5 Trp Gly Tyr Val 830	Glu  Ser  His  Leu  Ala  815  Leu	Ala Cys Ile Trp 800 Tyr	Asn Lys Pro Asp	Gly Phe Ala
Glu Phe Arg 785 Cys	Lys Leu 770 Val Leu Val	Leu 755 Tyr Val Leu His	Trp 40 Val Phe Pro Leu 80 Gly	Met Val Val Leu 79 Leu 5	Leu Ile 775 Ala O Ala Val	His 760 Phe Ala Leu	Leu 745 Ala Leu Tyr Pro Ala 825	Leu Ala Val Ser Gln 310 Ala	Ser Ala 7 Leu 795 Gln Leu	Ala 76 Ala 80 Thr Ala Leu	75 Ala 5 Trp Gly Tyr Val 830	Glu  Ser  His  Leu  Ala  815  Leu	Ala Cys Ile Trp 800 Tyr	Asn Lys Pro Asp	Gly Phe Ala

	850				85	5			8	60					
Ala	Pro	Ser	Met	Gln	Ala	Arg	Gly	Gly	Arg	Asp	Gly	Ile	Ile	Trp	Ala
865				87	0			i	875				880	ı	
Ala	Thr	Ile	Phe	Cys	Pro	Gly	Val	۷al	Phe	Asp	Ile	Thr	Lys	Trp	Leu
			88	35				890				895	5		
Leu	Ala	Val	Leu	Gly	Pro	Gly	Туг	Leu	Leu	Arg	Gly	Ala	Leu	Thr	Arg
		9	00				905				91	0			
Val	Pro	Tyr	Phe	Val	Arg	Ala	His	Ala	Leu	Leu	Arg	Met	Сув	Thr	Met
	9	15				920				92	25				
Val	Arg	His	Leu	Ala	Gly	Gly	Arg	Tyr	Val	Gln	Met	Ala	Leu	Leu	Ala
	930				935	5			9	40					
Leu	Gly	Arg	Trp	Thr	Gly	Thr	Tyr	Ile	Tyr	Asp	His	Leu	Thr	Pro	Met
945				95	0			9	955				960		
Ser	Asp	Trp	Ala	Ala	Ser	Gly	Leu	Arg	Asp	Leu	Ala	Val	Ala	Val	Glu
			96	5				970				975	i		
Pro	Ile	Ile	Phe	Ser	Pro	Met	Glu	Lys	ГЛЯ	Val	Ile	Val	Trp	Gly	Ala
		9	80				985				99	0			
Glu	Thr	Ala	Ala	Сув	Gly	Asp	Ile	Leu	His	Gly	Leu	Pro	Val	Ser	Ala
	9	95				1000				10	05				
Arg	Leu	Gly	Arg	Glu	Ile	Leu	Leu	Gly	Pro	Ala	qeA	Gly	Tyr	Thr	Ser
	010				101					20					
Lys	Gly	Trp	Lys	Leu	Leu	Ala	Pro	Ile	Thr	Ala	Tyr	Ala	Gln	Gln	Thr
1025				103					035				104		
Arg	Gly	Leu	Leu		Ser	Ile	Val	Val	Ser	Met	Thr	Gly	Arg	Asp	Lys
			104					050				105			
Thr	Glu		Ala	Gly	Glu			Val	Leu	Ser	Thr	Val	Thr	Gln	Ser
		10					1065				107				
Phe			Thr	Ser				Val	Leu			Val	Tyr	His	Gly
		175				1080				10					
		Asn	Lys	Thr			Gly	Ser			Pro	Val	Thr	Gln	Met
	090	_			1099					00					
Tyr	Ser	Ser	Ala	Glu	Gly	Asp	Leu	Val	Gly	Trp	Pro	Ser	Pro	Pro	Gly

110	5			11	10			1	115				112	0	
Thr	Lys	Ser	Leu	Glu	Pro	Сув	Thr	Сув	Gly	Ala	Val	Asp	Leu	Tyr	Leu
			112	25			1	130				113	5		
Va1	Thr	Arg	Asn	Ala	Asp	Val	Ile	Pro	Ala	Arg	Arg	Arg	Gly	Asp	Lys
		11	40			:	1145				115	50			
Arg	Gly	Ala	Leu	Leu	Ser	Pro	Arg	Pro	Leu	Ser	Thr	Leu	Lys	Gly	Ser
	1	155				1160	ı			11	65				
Ser	Gly	Gly	Pro	Va1	Leu	Cys	Pro	Arg	Gly	нів	Ala	Val	Gly	Ile	Phe
1	170				117	5			13	180					
Arg	Ala	Ala	Val	Суя	ser	Arg	Gly	Val	Ala	Lys	ser	Ile	Asp	Phe	Ile
1189	5			119	90			1	195				120	0	
Pro	Val	Glu	Thr	Leu	Asp	Ile	Val	Thr	Arg	Ser	Pro	Thr	Phe	Ser	Asp
			120	5			1.	210				121	5		
Asn	ser	Thr	Pro	Pro	Ala	Val	Pro	Gln	Thr	Tyr	Gln	Val	Gly	Tyr	Leu
		12	20			1	L 2 2 5				123	0			
His	Ala	Pro	Thr	Gly	Ser	Gly	Lys	Ser	Thr	Lys	Va 1	Pro	Va1	Ala	Tyr
	1:	235				1240				12	4 5				
Ala	Ala	Gln	Gly	Tyr	Lys	Val	Leu	Val	Leu	Asn	Pro	Ser	Val	Ala	Ala
1	250				125	5			12	260					
Thr	Leu	Gly	Phe	Gly	Ala	Tyr	Leu	Ser	Lys	Ala	нів	Gly	Ile	Asn	Pro
1265	i			127	0			1	275				128	כ	
Asn	Ile	Arg	Thr	Gly	Val	Arg	Thr	Val	Thr	Thr	Gly	Glu	Pro	Ile	Thr
			128	5			1	290				1295	ő		
ľyr	Ser	Thr	туг	Gly	ГЛЗ	Phe	Leu	Ala	Asp	Gly	Gly	Cys	Ala	Gly	Gly
		13	0 0			1	305				131	0			
Ala	Tyr	Asp	Ile	Ile	Ile	Cys	Asp	Glu	Cys	His	Ser	Val	Asp	Ala	Thr
	13	315				1320				13:	25				
Thr	Ile	Leu	Gly	Ile	Gly	Thr	Val	Leu	Asp	Gln	Ala	Glu	Thr	Ala	Gly
1.	330				1335	5			13	40					
/al	Arg	Leu	Thr	Val	Leu	Ala	Thr	Ala	Thr	Pro	Pro	Gly	Ser	Val	Thr
345				135	0			1:	355				1360	)	
Th so	Dwa	uia	Dago	3	T1.	a1	01	77 m 7	7. J -	<b>T</b>	<b>01</b> -	a2	<b>~1</b>	<b>01</b>	α1.

			13	65			1	370				137	5		
Ile	Pro	Phe	Tyr	Gly	Arg	Ala	Phe	Pro	Leu	Ser	Tyr	Ile	Lys	Gly	Gly
		13	80				1385				139	<b>∂</b> 0			
Arg	His	Leu	Ile	Phe	Cys	His	Ser	Lys	Lys	Lys	Сув	Asp	Glu	Leu	Ala
	1	395				1400	)			14	05				
Thr	Ala	Leu	Arg	Gly	Met	Gly	Leu	Asn	Ala	Val	Ala	Tyr	Tyr	Arg	Gly
1	410				141	5			1	420					
Leu	Asp	Va1	Ser	Ile	Ile	Pro	Thr	Gln	Gly	Asp	Val	Val	Val	Val	Ala
142	5			14	3 0			1	435				144	0	
Thr	Asp	Ala	Leu	Met	Thr	Gly	Tyr	Thr	Gly	Asp	Phe	Asp	Ser	Val	Ile
			144	<del>1</del> 5			1	450				145	5		
Asp	Сув	Asn	Val	Ala	Val	Thr	Gln	Ala	Val	Asp	Phe	Ser	Leu	Asp	Pro
		14	60			:	1465				147	0			
Thr	Phe	Thr	Ile	Thr	Thr	Gln	Thr	Val	Pro	Gln	Asp	Ala	Val	Ser	Arg
	14	175				1480				14	85				
Ser	Gln	Arg	Arg	Gly	Arg	Thr	Gly	Arg	Gly	Arg	Leu	Gly	Tle	Tyr	Arg
1	490				149	5			15	500					
Tyr	Val	Ser	Thr	Gly	Glu	Arg	Ala	Ser	Gly	Met	Phe	Asp	Ser	Val	Val
150	5			151	0			1	515				152	o	
Leu	Сув	Glu	Сув	Tyr	Asp	Ala	Gly	Ala	Ala	Trp	Tyr	Glu	Leu	Ser	Pro
			152	:5			1	530				1539	ő		
Val	Glu	Thr	Thr	Val	Arg	Leu	Arg	Ala	туг	Phe	Asn	Thr	Pro	Gly	Leu
		15	40			1	545				155	0			
Pro	Val	Сув	Gln	Asp	His	Leu	Glu	Phe	Trp	G1u	Ala	Val	Phe	Thr	Gly
	15	55				1560				15	55				
Leu	Thr	His	Ile	Asp	Ala	His	Phe	Leu	Ser	Gln	Thr	гуя	Gln	Ser	Gly
1	570				157	5			15	80					
Glu	Asn	Phe	Ala	Tyr	Leu	Val	Ala	Tyr	Gln	Ala	Thr	Val	Cys	Ala	Arg
1585	5			159	0			1	595				1600	)	
Ala	Lys	Ala	Pro	Pro	Pro	Ser	Trp	Asp	Val	Met	Trp	Ľуs	Сув	Leu	Thr
			160	5			1	610				1619	<b>i</b>		
λνα	T. 🔾 11	Liza	nro	mb v	Lan	3703	a1	D	m la sa	D	T	÷		_	•

	1620			1	L625				163	0					
Gly	Ser	Val	Thr	Asn	Glu	Val	Thr	Leu	Thr	His	Pro	Val	Thr	Lys	Tyr
	1:	635				1640				16	45				
Ile	Ala	Thr	Cys	Met	Gln	Ala	Asp	Leu	Glu	Val	Met	Thr	Ser	Thr	Trp
1	650				165	5			16	60					
Val	Leu	Ala	Gly	Gly	Val	Leu	Ala	Ala	Val	Ala	Ala	Tyr	Сув	Leu	Ala
1669	5			167	70			1	675				1680	ס	
Thr	Gly	Cys	Val	Ser	11e	Ile	Gly	Arg	Leu	His	Ile	Asn	Gln	Arg	Ala
			168	35			1.	690				169	5		
Val	Val	Ala	Pro	Asp	Lys	Glu	Va1	Leu	Tyr	Glu	Ala	Phe	Asp	G1u	Met
		17	00		•	1	1705				171	.0			
Glu	Glu	Cys	Ala	Ser	Arg	Ala	Ala	Leu	Leu	Glu	Glu	Gly	Gln	Arg	lle
	1	715				1720				17:	25				
Ala	Glu	Met	Leu	Lys	Ser	Lys	Ile	Gln	Gly	Leu	Leu	Gln	Gln	Ala	Ser
1	730				173	5			17	40					
Lys	Gln	Ala	Gln	Asp	Ile	Gln	Pro	Ala	۷al	Gln	Ala	Ser	Trp	Pro	Lys
1745	ō			175	i ()			1	755				1760		
		Gln	Phe			Lys	ніѕ			Asn	Phe	Ile			Ile
		Gln	Phe 176	ттр		Lys				Asn	Phe	Ile 1779	Ser		Ile
Met	Glu		176	Trp	Ala	Lys Ser	1	Меt 770	Trp			1779	Ser	Gly	
Met	Glu		176 Ala	Trp	Ala	Ser	1	Меt 770	Trp			1779 Pro	Ser	Gly	
Met Gln	Glu Tyr	Leu 17	176 Ala 80	Trp 55 Gly	Ala Leu	Ser	1 Thr .785	Met 770 Leu	Trp Pro	Gly	Asn 179	1779 Pro 0	ser 5 Ala	Gly Val	Ala
Met Gln	Glu Tyr Met	Leu 17	176 Ala 80	Trp 55 Gly	Ala Leu	Ser	1 Thr .785	Met 770 Leu	Trp Pro	Gly	Asn 179 Pro	1779 Pro 0	ser 5 Ala	Gly Val	Ala
Met Gln Ser	Glu Tyr Met	Leu 17 Met 795	176 Ala 80 Ala	Trp 55 Gly Phe	Ala Leu Ser	Ser J Ala	1 Thr .785 Ala	Met 770 Leu Leu	Trp Pro Thr	Gly Ser 180	Asn 179 Pro	1779 Pro 0 Leu	Ser 5 Ala Ser	Gly Val	Ala Ser
Met Gln Ser	Glu Tyr Met	Leu 17 Met 795	176 Ala 80 Ala	Trp 55 Gly Phe	Ala Leu Ser	Ser J Ala 1800 Ile	1 Thr .785 Ala	Met 770 Leu Leu	Trp Pro Thr	Gly Ser 180	Asn 179 Pro	1779 Pro 0 Leu	Ser 5 Ala Ser	Gly Val	Ala Ser
Met Gln Ser Thr	Glu Tyr Met 17 Thr 810	Leu 17 Met 795 Ile	176 Ala 80 Ala Leu	Trp 55 Gly Phe Leu	Ala Leu Ser Asn	Ser J Ala 1800 Ile	Thr .785 Ala Leu	Met 770 Leu Leu	Trp Pro Thr Gly	Gly Ser 180 Trp	Asn 179 Pro )5 Leu	1779 Pro 0 Leu Ala	Ser Ala Ser Ser	Gly Val Thr	Ala Ser Ile
Met Gln Ser Thr	Glu Tyr Met 17 Thr 810 Pro	Leu 17 Met 795 Ile	176 Ala 80 Ala Leu	Trp 55 Gly Phe Leu	Ala Leu Ser Asn 1819	Ser J Ala 1800 Ile	Thr .785 Ala Leu	Met 770 Leu Leu Gly	Trp Pro Thr Gly	Gly Ser 180 Trp	Asn 179 Pro )5 Leu	1779 Pro 0 Leu Ala	Ser Ala Ser Ser	Gly Val Thr Gln Val	Ala Ser Ile
Met Gln Ser Thr 1 Ala 1825	Tyr Met 17 Thr 810 Pro	Leu 17 Met 795 Ile	176 Ala 80 Ala Leu Ala	Trp 55 Gly Phe Leu Gly 183	Ala Leu Ser Asn 1819 Ala	Ser J Ala 1800 Ile	Thr .785 Ala Leu Gly	Met 770 Leu Leu Gly Phe	Trp Pro Thr Gly 18 Val	Gly Ser 180 Trp 20 Val	Asn 179 Pro )5 Leu Ser	1779 Pro 0 Leu Ala	Ser  Ala  Ser  Ser  Leu  1840	Gly Val Thr Gln Val	Ala Ser Ile Gly
Met Gln Ser Thr 1 Ala 1825	Tyr Met 17 Thr 810 Pro	Leu 17 Met 795 Ile	176 Ala 80 Ala Leu Ala	Trp 55 Gly Phe Leu Gly 183 Ser	Ala Leu Ser Asn 1819 Ala	Ser Ala 1800 Ile 5	Thr .785 Ala Leu Gly	Met 770 Leu Leu Gly Phe	Trp Pro Thr Gly 18 Val	Gly Ser 180 Trp 20 Val	Asn 179 Pro )5 Leu Ser	1779 Pro 0 Leu Ala	Ser  Ala  Ser  Ser  Leu  1840 Asp	Gly Val Thr Gln Val	Ala Ser Ile Gly
Met Gln Ser Thr 1 Ala 1825 Ala	Tyr  Met  17  Thr  810  Pro  Ala	Leu 17 Met 795 Ile Pro	176 Ala 80 Ala Leu Ala Gly 184	Trp  5  Gly  Phe  Leu  Gly  183  Ser	Ala Leu Ser Asn 1819 Ala 0	Ser Ala 1800 Ile 5	Thr .785 Ala Leu Gly Leu	Met 770 Leu Leu Gly Phe 1 Gly 850	Trp Pro Thr Gly 18 Val 835 Lys	Gly Ser 180 Trp 20 Val	Asn 179 Pro D5 Leu Ser Leu	Pro  Leu  Ala  Gly  Val	Ser Ala Ser Leu 1840 Asp	Gly Val Thr Gln Val	Ala Ser Ile Gly Leu
Met Gln Ser Thr 1 Ala 1825 Ala	Tyr  Met  17  Thr  810  Pro  Ala	Leu 17 Met 795 Ile Pro	176 Ala 80 Ala Leu Ala Gly 184	Trp  5  Gly  Phe  Leu  Gly  183  Ser	Ala Leu Ser Asn 1819 Ala 0	Ser  Ala 1800 Ile 5 Thr Gly	Thr .785 Ala Leu Gly Leu	Met 770 Leu Leu Gly Phe 1 Gly 850	Trp Pro Thr Gly 18 Val 835 Lys	Gly Ser 180 Trp 20 Val	Asn 179 Pro D5 Leu Ser Leu	1779 Pro  Leu  Ala  Gly  Val  1859 Ala	Ser Ala Ser Leu 1840 Asp	Gly Val Thr Gln Val	Ala Ser Ile Gly Leu

	1	375				1880				18	8 5				
Gly	Ile	Leu	Ser	Pro	Gly	Ala	Leu	Val	Val	Gly	Val	Ile	Сув	Ala	Ala
1	890				189	5			19	900					
Ile	Leu	Arg	Arg	His	Val	Gly	Pro	Gly	Glu	Gly	Ala	Val	Gln	Trp	Met
190	5			191	0			1	915				1920	)	
Asn	Arg	Leu	Ile	Ala	Phe	Ala	Ser	Arg	Gly	Asn	His	٧al	Ala	Pro	Thr
			192	25			1	930				193	5		
His	Tyr	Val	Thr	Glu	ser	Asp	Ala	Ser	Gln	Arg	Val	Thr	Gln	Leu	Leu
		19	40			:	1945				195	0			
Gly	Ser	Leu	Thr	Ile	Thr	Ser	Leu	Leu	Arg	Arg	Leu	His	Asn	Trp	Ile
	15	955				1960				19	65				
Thr	Glu	Asp	Сув	Pro	Ile	Pro	Cys	Ala	Gly	Ser	Trp	Leu	Arg	Asp	Val
1	970				197	5			19	089					
Trp	Asp	Trp	Val	Cys	Thr	Ile	Leu	Thr	Asp	Phe	Гуs	Äsn	Trp	Leu	Thr
198	5			199	0			1	995				2000	)	
Ser	Lys	Leu	Phe	Pro	Lys	Met	Pro	Gly	Leu	Pro	Phe	I1e	Ser	Cys	Gln
			200	) 5			2	010				2019	5		
Lys	Gly	Tyr	Lys	Gly	Val	Trp	Ala	Gly	Thr	Gly	Ile	Met	Thr	Thr	Arg
		20	20			2	2025				203	0			
Cys	Pro	Cys	Gly	Ala	Asn	Ile	Ser	Gly	Asn	Val	Arg	Leu	Gly	Ser	Met
	20	35				2040				20	4 5				
Arg	Ile	Thr	Gly	Pro	Lys	Thr	Сув	Met	Asn	Thr	Trp	Gln	Gly	Thr	Phe
2	050				205	5			20	60					
Pro	Ile	Asn	Cys	Tyr	Thr	Glu	Gly	Gln	Cys	Leu	Pro	гÀа	Pro	Ala	Leu
2065	5			207	0			2	075				2080	)	
Asn	Phe	Lys	Thr	Ala	Ile	Trp	Arg	Val	Ala	Ala	Ser	Glu	Tyr	Ala	Glu
			208	5			2	090				2095	5		
Val	Thr	Gln	His	Gly	Ser	Tyr	Ala	Tyr	Ile	Thr	Gly	Leu	Thr	Thr	Asp
		21	00			2	105				211	0			
Asn	Leu	Гув	Val	Pro	Cys	Gln	Leu	Pro	Ser	Pro	Glu	Phe	Phe	Ser	Trp
	21	15				2120				21:	25				
	21									41.					

2	130				213	5			21	4 0					
Phe	Arg	Asp	Glu	Val	Ser	Phe	Ser	Val	Gly	Leu	Asn	Ser	Phe	Val	Val
2145	วี			215	0			2	155				2160	)	
Gly	Ser	Gln	Leu	Pro	Сув	Asp	Pro	Glu	Pro	Asp	Thr	Glu	Val	Val	Met
			216	5			2	170				2175	Š		
Ser	Met	Leu	Thr	Asp	Pro	Ser	His	Ile	Thr	Ala	Glu	Ala	Ala	Ala	Arg
		21	80			2	185				219	0			
Arg	Leu	Ala	Arg	Gly	Ser	Pro	Pro	Ser	Glu	Ala	ser	Ser	Ser	Ala	Ser
	21	L95				2200				22	05				
Gln	Leu	Ser	Ala	Pro	Ser	Leu	Arg	Ala	Thr	Сув	Thr	Thr	His	Gly	Arg
2	210				221	5			22	20					
Thr	Tyr	Asp	Val	Asp	Met	Val	Asp	Ala	Asn	Leu	Phe	Met	Gly	Gly	Gly
2225	5			223	0			2	235				2240	0	
Va1	Ile	Arg	Ile	Glu	Ser	Glu	Ser	ГЛЗ	Val	Val	Val	Leu	Asp	Ser	Leu
			224	5			2	250				225	5		
Asp	Ser	Met	Thr	Glu	Glu	Glu	Gly	Asp	Leu	Glu	pro	Ser	Val	Pro	Ser
		22	60			2	2265				227	0			
G1 u	Tyr	Met	Leu	Pro	Arg	Lys	Arg	Phe	Pro	Pro	Ala	Leu	Pro	Ala	Trp
	22	275				2280				22	8 5				
Ala	Arg	Pro	Asp	Tyr	Asn	Pro	Pro	Leu	Val	Glu	Ser	Trp	Lys	Arg	Pro
2	290				229	5			23	00					
Asp	Tyr	Gln	Pro	Pro	Thr	Val	Ala	Gly	Cys	Ala	Leu	Pro	Pro	Pro	Lys
2305	5			231	L <b>0</b>			2	315				2320	0	
Lys	Thr	Pro	Thr	Pro	Pro	Pro	Arg	Arg	Arg	Arg	Thr	Val	Gly	Leu	Ser
			232	:5			2	330				233	5		
Glu	ser	Thr	Ile	Gly	Asp	Ala	Leu	Gln	Gln	Leu	Ala	Ile	Lys	Ser	Phe
		23					2345				235				
Gly	Gln	Pro	Pro	Pro	Ser	Gly	Asp	Ser	Gly	Leu	Ser	Thr	Gly	Ala	Asp
-		355				2360				23					
Ala			Ser	Gly	Asp	Arg	Thr	Pro	Pro	Asp	Glu	Leu	Ala	Leu	Ser
	370	•		•	237					80					
		Gly	Sar	ጥ <sub>ኮ ድ</sub>			Met	Pro	Pro	Len	Glu	alv	Glu	Pro	Glv

2385	5			239	0			2	395				2400	)	
qsA	Pro	Asp	Leu	Glu	Pro	Glu	Gln	Val	Glu	Leu	Gln	Pro	Pro	Pro	Gln
			240	) 5			2	410				2415	5		
Gly	Gly	Glu	Ala	Ala	Pro	Gly	Ser	Asp	Ser	Gly	Ser	Trp	ser	Thr	Сув
		24	20			2	2425				243	0			
Ser	Glu	Glu	Asp	Asp	Ser	Val	Val	Cys	Сув	Ser	Met	Ser	Tyr	ser	Trp
	24	435				2440				24	45				
Thr	Gly	Ala	Leu	Ile	Thr	Pro	Сув	Ser	Pro	Glu	Glu	Glu	Lув	Leu	Pro
2	450				245	5			24	60					
Ile	Asn	Ser	Leu	Ser	Asn	Ser	Leu	Leu	Arg	туг	His	Asn	Lys	Val	Tyr
2469	5			247	0			2	475				2480	)	
Cys	Thr	Thr	Ser	Lys	Ser	Ala	ser	Leu	Arg	Ala	ьув	Lys	Val	Thr	Phe
			248	5			2	490				249	5		
Asp	Arg	Met	Gln	Val	Leu	Asp	Ala	Tyr	Tyr	Asp	Ser	Val	Leu	Lys	Asp
		25	00			2	2505		•		251	0			
Ile	Lys	Leu	Ala	Ala	Ser	Lys	Va1	Ser	Ala	Arg	Leu	Leu	Thr	Leu	Glu
	25	515				2520				25	25				
Glu	Ala	Сув	Gln	Leu	Thr	Pro	Pro	His	Ser	Ala	Arg	Ser	Lys	Tyr	Gly
2	530				253	5			25	340					
Phe	Gly	Ala	Lys	Glu	Val	Arg	Ser	Leu	Ser	Gly	Arg	Ala	Val	Asn	His
2545	5			255	0			2	555				2560	)	
Ile	Lys	Ser	Val	Trp	ьув	Asp	Ľеu	Leu	Glu	Asp	Ser	Gln	Thr	Pro	Ile
			256	5 5			2	570				2579	5		
Pro	Thr	Thr	Ile	Met	Ala	Lys	Asn	Glu	Val	Phe	Сув	۷al	Asp	Pro	Ala
		25	80			2	8585				259	0			
Lys	Gly	Gly	Lys	Lys	Pro	Ala	Arg	Leu	Ile	Val	Tyr	Pro	Asp	Leu	Gly
	25	595				2600				26	05				
Val	Arg	Val	Сув	Glu	Lys	Met	Ala	Leu	Tyr	Asp	Val	Thr	Gln	Lys	Leu
2	610				261	5			26	20					
Pro	Gln	Ala	Val	Met	Gly	Ala	Ser	туг	Gly	Phe	Gln	Tyr	Ser	Pro	Ala
2625	5			263	0			2	635				2640	)	
											Glu				

			264	ł 5			2	650				265	5		
Met	Gly	Phe	Ser	Tyr	Asp	Thr	Arg	Cys	Phe	Asp	ser	Thr	Val	Thr	Glu
		26	60			:	2665				267	0			
Arg	Asp	Ile	Arg	Thr	Glu	Glu	Ser	Ile	Tyr	Gln	Ala	Суз	Ser	Leu	Pro
	26	575				2680				26	85				
Glu	Glu	Ala	Arg	Thr	Ala	Ile	His	Ser	Leu	Thr	Glu	Arg	Leu	Tyr	Val
2	690				269	5			27	00					
Gly	Gly	Pro	Met	Phe	Asn	Ser	ьуs	Gly	Gln	Ser	Cys	Gly	Tyr	Arg	Arg
2709	5			273	0			2	715				2720	)	
Сув	Arg	Ala	Ser	Gly	Val	Leu	Thr	Thr	Ser	Met	Gly	Asn	Thr	Ile	Thr
			272	25			2	730				273	5		
Cys	Tyr	Val	Lys	Ala	Leu	Ala	Ala	Сув	ГАв	Ala	Ala	Gly	Ile	Ile	Ala
		27	40			2	2745				275	0			
Pro	Thr	Met	Leu	Val	Cys	Gly	Asp	Asp	Leu	Val	Val	Ile	Ser	Glu	Ser
	27	755				2760				27	65				
Gln	Gly	Thr	Glu	Glu	Asp	Glu	Arg	Asn	Leu	Arg	Ala	Phe	Thr	Glu	Ala
2	770				277	5			27	80	٠				
		Arg	туг	Ser			Pro	Gly			Pro	Arg	Pro	Glu	туг
	Thr	Arg	Туг	Ser 279	Ala		Pro	_			Pro	Arg	Pro 2800		Туг
Met 2785	Thr 5		Tyr	279	Ala	Pro		2	Asp 795	Pro			2800	)	
Met 2785	Thr 5			279 Ile	Ala	Pro	Сув	2	Asp 795	Pro			2800 Val	o	
Met 2785 Asp	Thr 5 Leu	Glu	Leu	279 Ile )5	Ala 00 Thr	Pro Ser	Cys 2	2 Ser 810	Asp 795 Ser	Pro	Val	Ser 281	2800 Val	) Ala	Leu
Met 2785 Asp	Thr 5 Leu	Glu	Leu 280 Gly	279 Ile )5	Ala 00 Thr	Pro Ser Arg	Cys 2	2 Ser 810	Asp 795 Ser	Pro	Val	Ser 2819 Asp	2800 Val	) Ala	Leu
Met 2785 Asp Gly	Thr Leu Pro	Glu Gln 28	Leu 280 Gly	279 Ile 05 Arg	Ala 00 Thr Arg	Pro Ser Arg	Cys 2 Tyr 2825	2 Ser 810 Tyr	Asp 795 Ser Leu	Pro Asn Thr	Val Arg 283	Ser 2819 Asp	2800 Val 5 Pro	) Ala Thr	Leu
Met 2785 Asp Gly	Thr Leu Pro	Glu Gln 28	Leu 280 Gly 20	279 Ile 05 Arg	Ala O Thr Arg	Pro Ser Arg	Cys 2 Tyr 2825 Glu	2 Ser 810 Tyr	Asp 795 Ser Leu	Pro Asn Thr	Val Arg 283 His	Ser 2819 Asp	2800 Val 5 Pro	) Ala Thr	Leu
Met 2785 Asp Gly Ser	Thr Leu Pro	Glu Gln 28 Ala 335	Leu 280 Gly 20	279 Ile 05 Arg Ala	Ala 00 Thr Arg	Pro Ser Arg Trp	Cys 2 Tyr 2825 Glu	2 Ser 810 Tyr	Asp 795 Ser Leu Val	Pro Asn Thr Arg	Val Arg 283 His	Ser 2819 Asp 0 Ser	2800 Val 5 Pro	Ala Thr	Leu Thr Asn
Met 2785 Asp Gly Ser	Thr Leu Pro	Glu Gln 28 Ala 335	Leu 280 Gly 20 Arg	279 Ile 05 Arg Ala	Ala 00 Thr Arg	Pro Ser Arg Trp 2840	Cys 2 Tyr 2825 Glu	2 Ser 810 Tyr	Asp 795 Ser Leu Val	Pro Asn Thr Arg	Val Arg 283 His	Ser 2819 Asp 0 Ser	2800 Val 5 Pro	Ala Thr	Leu Thr Asn
Met 2785 Asp Gly Ser Ser	Thr Leu Pro Ile 28 Trp 850	Glu Gln 28 Ala 335 Leu	Leu 280 Gly 20 Arg	279 Ile 05 Arg Ala Asn	Ala O Thr Arg Ala Ile 285	Pro Ser Arg Trp 2840 Ile	Cys 2 Tyr 2825 Glu Gln	ser 810 Tyr Thr	Asp 795 Ser Leu Val Ala	Pro Asn Thr Arg 28- Pro	Val Arg 283 His 45	Ser 2819 Asp 0 Ser	2800 Val 5 Pro Pro	Ala Thr Val	Leu Thr Asn Arg
Met 2785 Asp Gly Ser Ser	Thr Leu Pro Ile 28 Trp 850 Val	Glu Gln 28 Ala 335 Leu	Leu 280 Gly 20 Arg	279 Ile 05 Arg Ala Asn	Ala O Thr Arg Ala Ile 285:	Pro Ser Arg Trp 2840 Ile	Cys 2 Tyr 2825 Glu Gln	Ser 810 Tyr Thr Tyr	Asp 795 Ser Leu Val Ala	Pro Asn Thr Arg 28- Pro	Val Arg 283 His 45	Ser 2819 Asp 0 Ser	2800 Val 5 Pro Pro	Ala Thr Val Val	Leu Thr Asn Arg
Met 2788 Asp Gly Ser 2 Met 2865	Thr Leu Pro Ile 28 Trp 850 Val	Glu Gln 28 Ala 335 Leu Leu	Leu 280 Gly 20 Arg	279 Ile 5 Arg Ala Asn Thr	Ala O Thr Arg Ala Ile 285: His	Pro Ser Arg Trp 2840 Ile 5	Cys 2 Tyr 2825 Glu Gln Phe	2 Ser 810 Tyr Thr Tyr	Asp 795 Ser Leu Val Ala 28 Ile 875	Pro Asn Thr Arg 286 Pro 60 Leu	Val Arg 283 His 45 Thr	Ser 2819 0 Ser Ile	2800 Val 5 Pro Pro Trp Gln 2880	Ala Thr Val Val	Leu Thr Asn Arg
Met 2788 Asp Gly Ser 2 Met 2865	Thr Leu Pro Ile 28 Trp 850 Val	Glu Gln 28 Ala 335 Leu Leu	Leu 280 Gly 20 Arg Gly Met	279 Ile 5 Arg Ala Asn Thr 287	Ala O Thr Arg Ala Ile 285: His	Pro Ser Arg Trp 2840 Ile 5	Cys 2 Tyr 2825 Glu Gln Phe	2 Ser 810 Tyr Thr Tyr	Asp 795 Ser Leu Val Ala 28 Ile 875	Pro Asn Thr Arg 286 Pro 60 Leu	Val Arg 283 His 45 Thr	Ser 2819 0 Ser Ile	2800 Val  Pro  Trp  Gln  2880  Tyr	Ala Thr Val Val	Leu Thr Asn Arg

2900 2905 2910

Ala Phe Ser Leu His Thr Tyr Thr Pro His Glu Leu Thr Arg Val Ala

2915 2920 2925

Ser Ala Leu Arg Lys Leu Gly Ala Pro Pro Leu Arg Ala Trp Lys Ser

2930 2935 2940

Arg Ala Arg Ala Val Arg Ala Ser Leu Ile Ser Arg Gly Gly Arg Ala

2945 2950 2955 2960

Ala Val Cys Gly Arg Tyr Leu Phe Asn Trp Ala Val Lys Thr Lys Leu

2965 2970 2975

Lys Leu Thr Pro Leu Pro Glu Ala Arg Leu Leu Asp Leu Ser Ser Trp

2980 2985 2990

Phe Thr Val Gly Ala Gly Gly Gly Asp Ile Tyr His Ser Val Ser Arg

2995 3000 3005

Ala Arg Pro Arg Leu Leu Leu Ser Leu Leu Leu Leu Ser Val Gly

3010 3015 3020

Val Gly Leu Phe Leu Leu Pro Ala Arg

3025 3030

<210> 7

<211> 8024

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: replicon

<400> 7

accugecceu aauaggggeg acaeueegee augaaueaeu eeeeugugag gaacuaeugu 60 euucaegeag aaagegeeua gecauggegu uaguaugagu gueguaeage euceaggeee 120 eeeeeueeeg ggagageeau aguggueuge ggaaceggug aguaeaeegg aauugeeggg 180

aagacugggu	ccuuucuugg	auaaacccac	ucuaugeceg	gccauuuggg	cgugcccccg	240
caagacugcu	agccgaguag	cguuggguug	cgaaaggccu	ugugguacug	ccugauaggg	300
cgcuugcgag	ugeceeggga	ggucucguag	accgugcacc	augagcacaa	auccuaaacc	360
ucaaagaaaa	accaaaagaa	acaccaaccg	ucgcccaaug	auugaacaag	auggauugca	420
cgcagguucu	ccggccgcuu	ggguggagag	gcuauucggc	uaugacuggg	cacaacagac	480
aaucggcugc	ucugaugeeg	ccguguuccg	gcugucagcg	caggggcgcc	cgguucuuuu	540
ugucaagacc	gaccuguccg	gugcccugaa	ugaacugcag	gacgaggcag	cgcggcuauc	600
guggcuggcc	acgacgggcg	uuccuugcgc	agcugugcuc	gacguuguca	cugaagcggg	660
aagggacugg	cugcuauugg	gcgaagugcc	ggggcaggau	cuccugucau	cucaccuugc	720
uccugocgag	aaaguaucca	ucauggcuga	ugcaaugcgg	cggcugcaua	cgcuugaucc	780
ggcuaccugc	ccauucgacc	accaagcgaa	acaucgcauc	gagcgagcac	guacucggau	840
ggaagccggu	cuugucgauc	aggaugaucu	ggacgaagag	caucaggggc	ucgegecage	900
cgaacuguuc	gccaggcuca	aggcgcgcau	gcccgacggc	gaggaucucg	ucgugaccca	960
uggegaugee	ugcuugccga	auaucauggu	ggaaaauggc	cgcuuuucug	gauucaucga	1020
cuguggccgg	cugggugugg	cggaccgcua	ucaggacaua	geguuggeua	cccgugauau	1080
ugcugaagag	cuuggcggcg	aaugggcuga	cegeuuccuc	gugcuuuacg	guaucgeege	1140
ucccgauucg	cagegeaueg	ccuucuaucg	ccuucuugac	gaguucuucu	gaguuuaaac	1200
ccucucccuc	ccccccccu	aacguuacug	gccgaagccg	cuuggaauaa	ggccggugug	1260
cguuugucua	uauguuauuu	uccaccauau	ugccgucuuu	uggcaaugug	agggcccgga	1320
aaccuggccc	ugucuucuug	acgagcauuc	cuaggggucu	uuccccucuc	gccaaaggaa	1380
ugcaaggucu	guugaauguc	gugaaggaag	caguuccucu	ggaagcuucu	ugaagacaaa	1440
caacgucugu	agcgacccuu	ugcaggcagc	ggaacccccc	accuggcgac	aggugccucu	1500
gcggccaaaa	gccacgugua	uaagauacac	cugcaaaggc	ggcacaaccc	cagugecacg	1560
uugugaguug	gauaguugug	gaaagaguca	aauggcucuc	cucaagcgua	uucaacaagg	1620
ggcugaagga	ugcccagaag	guaccccauu	guaugggauc	ugaucugggg	ccucggugca	1680
caugcuuuac	auguguuuag	ucgagguuaa	aaaaacgucu	aggccccccg	aaccacgggg	1740
acgugguuuu	ccuuugaaaa	acacgaugau	accauggcuc	ccaucacugc	uuaugcccag	1800
caaacacgag	gccuccuggg	cgccauagug	gugaguauga	cggggcguga	caggacagaa	1860
caggccgggg	aaguccaaau	ccuguccaca	gucucucagu	ccuuccucgg	aacaaccauc	1920
ucggggguuu	uguggacugu	uuaccacgga	gcuggcaaca	agacucuagc	cggcuuacgg	1980
gguccgguca	cgcagaugua	cucgagugcu	gagggggacu	ugguaggcug	geceagecee	2040
ccugggacca	agucuuugga	gccgugcaag	uguggagccg	ucgaccuaua	ucuggucacg	2100

cggaacgcug	augucauccc	ggcucggaga	cgcggggaca	agcggggagc	auugcucucc	2160
ccgagaccca	uuucgaccuu	gaaggggucc	ucgggggggc	eggugeueug	cccuaggggc	2220
cacgucguug	ggcucuuccg	agcagcugug	ugcucucggg	gcguggccaa	auccaucgau	2280
uucauccccg	uugagacacu	cgacguuguu	acaaggucuc	ccacuuucag	ugacaacagc	2340
acgccaccgg	cugugeecea	gaccuaucag	gueggguaeu	ugcaugeuce	aacuggcagu	2400
ggaaagagca	ccaagguccc	ugucgcguau	geegeeeagg	gguacaaagu	acuagugcuu	2460
aaccccucgg	uagcugccac	ccugggguuu	ggggcguacc	uauccaaggc	acauggcauc	2520
aaucecaaca	uuaggacugg	agucaggacc	gugaugaccg	gggaggccau	cacguacucc	2580
acauauggca	aauuucucgc	cgaugggggc	ugcgcuagcg	gcgccuauga	caucaucaua	2640
ugcgaugaau	gecaegeúgu	ggaugeuace	uccauucucg	gcaucggaac	gguccuugau	2700
caagcagaga	cagccggggu	cagacuaacu	gugcuggcua	cggccacacc	occeggguca	2760
gugacaaccc	cccaucccga	uauagaaaag	guaggccucg	ggcgggaggg	ugagaucccc	2820
uucuauggga	gggegauuce	ccuauccuge	aucaagggag	ggagacaccu	gauuuucugc	2880
cacucaaaga	aaaaguguga	cgagcucgcg	geggeeeuue	ggggcauggg	cuugaaugcc	2940
guggcauacu	auagaggguu	ggacgueuce	auaauaccag	cucagggaga	uguggugguc	3000
gucgccaccg	acgeceucau	gacgggguac	acuggagacu	uugacuccgu	gaucgacugc	3060
aauguagcgg	ucacccaage	ugucgacuuc	agccuggacc	ccaccuucac	uauaaccaca	3120
cagacugucc	cacaagaege	ugucucacgc	agucagogoc	gcgggcgcac	agguagagga	3180
agacagggca	cuuauaggua	uguuuccacu	ggugaacgag	ccucaggaau	guuugacagu	3240
guagugcuuu	gugagugcua	cgacgcaggg	gcugcguggu	acgaucucac	accageggag	3300
accaccguca	ggcuuagagc	guauuucaac	acgcccggcc	uacccgugug	ucaagaccau	3360
cuugaauuuu	gggaggcagu	uuucaccggc	cucacacaca	uagacgccca	cuuccucuce	3420
caaacaaagc	aagcggggga	gaacuucgcg	uaccuaguag	ccuaccaagc	uacggugugc	3480
gccagagcca	aggeeceuce	cccguccugg	gacgccaugu	ggaagugccu	ggecegaeue	3540
aagccuacgc	uugegggeee	cacaccucuc	cuguaccguu	ugggcccuau	uaccaaugag	3600
gucacccuca	cacacccugg	gacgaaguac	aucgccacau	gcaugcaagc	ugaccuugag	3660
gucaugacca	gcacgugggu	ccuagcugga	ggaguccugg	cagccgucgc	cgcauauugc	3720
cuggcgacug	gaugeguuue	caucaucggc	cgcuugcacg	ucaaccagcg	agucgucguu	3780
gcgccggaua	aggagguecu	guaugaggcu	uuugaugaga	uggaggaaug	cgccucuagg	3840
gcggcucuca	ucgaagaggg	gcagcggaua	gccgagaugu	ugaaguccaa	gauccaaggc	3900
uugcugcagc	aggecucuaa	gcaggcccag	gacauacaac	ccgcuaugca	ggcuucaugg	3960
cccaaagugg	aacaauuuug	ggccagacac	auguggaacu	ucauuagegg	cauccaauac	4020

cucgcaggau	ugucaacacu	gccagggaac	cccgcggugg	cuuccaugau	ggcauucagu	4080
gccgcccuca	ccaguccguu	gucgaccagu	accaccaucc	uucucaacau	caugggaggc	4140
ugguuagcgu	cccagaucgc	accacccgcg	ggggccaccg	gcuuugucgu	caguggccug	4200
gugggggcug	ccgugggcag	cauaggccug	gguaaggugc	ugguggacau	ccuggcagga	4260
uauggugcgg	gcauuucggg	ggcccucguc	gcauucaaga	ucaugucugg	cgagaagccc	4320
ucuauggaag	augucaucaa	ucuacugccu	gggauccugu	cuccgggagc	ccugguggug	4380
ggggucaucu	gegeggeeau	ucugegeege	cacgugggac	cgggggaggg	egegguecaa	4440
uggaugaaca	ggcuuauugc	cuuugcuucc	agaggaaacc	acguegeece	uacucacuac	4500
gugacggagu	cggaugcguc	gcagcgugug	acccaacuac	uuggcucucu	uacuauaacc	4560
agccuacuca	gaagacucca	caauuggaua	acugaggacu	gececauece	augcuccgga	4620
uccuggcucc	gcgacgugug	ggacuggguu	ugcaccaucu	ugacagacuu	caaaaauugg	4680
cugaccucua	aauuguuccc	caagcugeee	ggccuccccu	ucaucucuug	ucaaaagggg	4740
uacaagggug	ugugggccgg	cacuggcauc	augaccacge	geugeeeuug	eggegecaac	4800
aucucuggca	auguccgccu	gggcucuaug	aggaucacag	ggccuaaaac	cugcaugaac	4860
accuggcagg	ggaccuuucc	uaucaauugc	uacacggagg	gccagugcgc	gccgaaaccc	4920
cccacgaacu	acaagaccgc	caucuggagg	guggeggeeu	cggaguacgc	ggaggugacg	4980
cagcaugggu	cguacuccua	uguaacagga	cugaccacug	acaaucugaa	aauuccuugc	5040
caacuaccuu	cuccagaguu	uuucuccugg	guggacggug	ugcagaucca	uagguuugca	5100
cccacaccaa	agccguuuuu	ccgggaugag	gucucguucu	gcguugggcu	uaauuccuau	5160
gcugucgggu	cccagcuucc	cugugaaccu	gagecegaeg	cagacguauu	gagguccaug	5220
cuaacagauc	cgccccacau	cacggcggag	acugeggege	ggcgcuuggc	acggggauca	5280
ccuccaucug	aggegageue	cucagugage	cagcuaucag	caccgucgcu	gegggeeace	5340
ugcaccaccc	acagcaacac	cuaugacgug	gacauggucg	augecaaceu	gcucauggag	5400
ggcggugugg	cucagacaga	gccugagucc	agggugcccg	uucuggacuu	ucucgageca	5460
auggccgagg	aagagagega	ceuugageee	ucaauaccau	cggagugcau	gcuccccagg	5520
agcggguuuc	cacgggccuu	accggcuugg	gcacggccug	acuacaacce	gccgcucgug	5580
gaaucgugga	ggaggccaga	uuaccaaccg	cccaccguug	cugguugugc	ucucccccc	5640
cccaagaagg	ccccgacgcc	ucccccaagg	agacgccgga	cagugggucu	gagcgagagc	5700
accauaucag	aageccucca	gcaacuggcc	aucaagaccu	uuggccagcc	ccccucgagc	5760
ggugaugeag	geueguecae	gggggcgggc	gccgccgaau	ccggcggucc	gacguccccu	5820
ggugageegg	cccccucaga	gacagguucc	gecuccucua	ugecececcu	cgagggggag	5880
ccuggagauc	cggaccugga	gucugaucag	guagagcuuc	aaccuccccc	ccaggggggg	5940

gggguagcuc	ccgguucggg	cucggggucu	uggucuacuu	gcuccgagga	ggacgauacc	6000
accgugugcu	gcuccauguc	auacuccugg	accggggcuc	uaauaacucc	cuguagecee	6060
gaagaggaaa	aguugccaau	caacccuuug	aguaacucgc	uguugcgaua	ccauaacaag	6120
guguacugua	caacaucaaa	gagegeeuca	cagagggcua	aaaagguaac	uuuugacagg	6180
acgcaagugc	ucgacgccca	uuaugacuca	gucuuaaagg	acaucaagcu	ageggeuuce	6240
aaggucagcg	caaggcuccu	caccuuggag	gaggegugee	aguugacucc	accccauucu	6300
gcaagaucca	aguauggauu	cggggccaag	gagguccgca	gcuuguccgg	gagggccguu	6360
aaccacauca	aguccgugug	gaaggaccuc	cuggaagacc	cacaaacacc	aauucccaca	6420
accaucaugg	ccaaaaauga	gguguucugc	guggaccccg	ccaagggggg	uaagaaacca	6480
gcucgccuca	ucguuuaccc	ugaccucggc	guccgggucu	gcgagaaaau	ggcccucuau	6540
gacauuacac	aaaagcuucc	ucaggcggua	augggagcuu	ccuauggcuu	ccaguacucc	6600
ccugcccaac	ggguggagua	ucucuugaaa	gcaugggcgg	aaaagaagga	ccccaugggu	6660
uuuucguaug	auacccgaug	cuucgacuca	accgucacug	agagagacau	caggaccgag	6720
gaguccauau	accaggccug	cucccugece	gaggaggccc	gcacugccau	acacucgcug	6780
acugagagac	uuuacguagg	agggcccaug	uucaacagca	agggucaaac	cugcgguuac	6840
agacguugcc	gcgccagcgg	ggugcuaacc	acuagcaugg	guaacaccau	cacaugcuau	6900
gugaaagccc	uagcggccug	caaggcugcg	gggauaguug	cgcccacaau	gcugguaugc	6960
ggcaaugacc	uaguagucau	cucagaaagc	caggggacug	aggaggacga	geggaaccug	7020
agagecuuca	cggaggccau	gaccagguac	ucugececue	cuggugauce	ccccagaccg	7080
gaaŭaugacc	uggagcuaau	aacauccugu	uccucaaaug	ugucuguggc	guugggcccg	7140
eggggeegee	gcagauacua	ccugaccaga	gacccaacca	cuccacucge	ccgggcugcc	7200
ugggaaacag	uuagacacuc	cccuaucaau	ucauggeugg	gaaacaucau	ccaguaugcu	7260
ccaaccauau	ggguucgcau	gguccuaaug	acacacuucu	ucuccauucu	caugguccaa	7320
gacacccugg	accagaaccu	caacuuugag	auguauggau	caguauacuc	cgugaauccu	7380
uuggaccuuc	cagccauaau	ugagagguua	cacgggcuug	acgccuuuuc	uaugcacaca	7440
uacucucacc	acgaacugac	geggguggeu	ucageceuca	gaaaacuugg	ggcgccaccc	7500
cucagggugu	ggaagagucg	ggcucgcgca	gucagggcgu	cccucaucuc	ccguggaggg	7560
aaagcggccg	uuugcggccg	auaucucuuc	aauugggcgg	ugaagaccaa	gcucaaacuc	7620
acuccauugc	cggaggcgcg	ccuacuggac	uuauccaguu	gguucaccgu	cggcgccggc	7680
gggggegaca	uuuuucacag	cgugucgcgc	gcccgacccc	gcucauuacu	cuucggccua	7740
cuccuacuuu	ucguaggggu	aggccucuuc	cuacuccccg	cucgguagag	cggcacacac	7800
uagguacacu	ccauagcuaa	cuguuccuuu	uuuuuuuuu	սսսսսսսսսս	นนนนนนนนนน	7860

uuucuuggug geuccaucuu ageecuague aeggeuageu gugaaaggue egugageege 7980 augacugeag agagugeegu aacuggueu ucugeagaue augu 8024

<210> 8

<211> 7994

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: replicon

<400> 8

accugecceu aauaggggeg acaeucegee augaaucaeu eeeeugugag gaacuaeugu 60 cuucacgcag aaagegeeua geeauggegu uaguaugagu gueguacage cuccaggeee 120 cccccucceg ggagagecau aguggucuge ggaaceggug aguacacegg aauugeeggg 180 aagacugggu ccuuucuugg auaaacccac ucuaugcccg gccauuuggg cgugcccccg 240 caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300 egeungegag ugeeeeggga gguenegnag accgugeace augageacaa aucenaaace 360 ucaaagaaaa accaaaagaa acaccaaccg ucgcccaaug auugaacaag auggauugca 420 cgcagguucu ccggccgcuu ggguggagag gcuauucggc uaugacuggg cacaacagac 480 aaucggeuge ueugaugeeg eeguguueeg geugueageg eaggggegee egguueuuuu 540 ugucaagacc gaccugucog gugcccugaa ugaacugcag gacgaggcag cgcggcuauc 600 guggeuggee acgaeggeg uuccuugege ageugugeue gaeguuguea eugaageggg 660 aagggacugg cugcuauugg gcgaagugcc ggggcaggau cuccugucau cucaccuugc 720 uccugcegag aaaguaucca ucauggcuga ugcaaugegg eggeugeaua egeuugauee 780 ggeuaceuge ecauucgace accaagegaa acaucgeaue gagegageae guacucggau 840 ggaagceggu cuuguegaue aggaugaueu ggaegaagag caucagggge uegegeeage 900 cgaacuguuc gccaggcuca aggcgcgcau gcccgacggc gaggaucucg ucgugaccca 960 uggegaugee ugeuugeega auaucauggu ggaaaaugge egeuuuueug gauueauega 1020 cuguggccgg cugggugugg cggaccgcua ucaggacaua gcguuggcua cccgugauau 1080

ugcugaagag	cuuggeggeg	aaugggcuga	ccgcuuccuc	gugcuuuacg	guaucgccgc	1140
ucccgauucg	cagegeaueg	ccuucuaucg	ccuucuugac	gaguucuucu	gaguuuaaac	1200
ccucucccuc	ccccccccu	aacguuacug	gccgaagccg	cuuggaauaa	ggccggugug	1260
cguuugucua	uauguuauuu	uccaccauau	ugccgucuuu	uggcaaugug	agggcccgga	1320
aaccuggccc	ugucuucuug	acgagcauuc	cuaggggucu	uuccccucuc	gccaaaggaa	1380
ugcaaggucu	guugaauguc	gugaaggaag	caguuccucu	ggaagcuucu	ugaagacaaa	1440
caacgucugu	agcgacccuu	ugcaggcagc	ggaacccccc	accuggcgac	aggugccucu	1500
gcggccaaaa	gccacgugua	uaagauacac	cugcaaaggc	ggcacaaccc	cagugecaeg	1560
uugugaguug	gauaguugug	gaaagaguca	aauggcucuc	cucaagcgua	uucaacaagg	1620
ggcugaagga	ugcccagaag	guaccccauu	guaugggauc	ugaucugggg	ccucggugca	1680
caugcuuuac	auguguuuag	ucgagguuaa	aaaaacgucu	aggececeeg	aaccacgggg	1740
acgugguuuu	ccuuugaaaa	acacgaugau	accauggcuc	ccaucacuge	uuaugeecag	1800
caaacacgag	gccuccuggg	cgccauagug	gugaguauga	cggggcguga	caggacagaa	1860
caggccgggg	aaguccaaau	ccuguccaca	gucucucagu	ccuuccucgg	aacaaccauc	1920
ucggggguuu	uguggacugu	uuaccacgga	gcuggcaaca	agacucuage	cggcuuacgg	1980
gguccgguca	cgcagaugua	cucgagugcu	gagggggacu	ugguaggcug	gcccagcccc	2040
ccugggacca	agucuuugga	gccgugcaag	uguggagccg	ucgaccuaua	ucuggucacg	2100
cggaacgcug	augucauccc	ggcucggaga	cgcggggaca	agcggggagc	auugcucucc	2160
ccgagaccca	uuucgaccuu	gaaggggucc	ncadaaaaac	eggugeueug	cccuaggggc	2220
cacgucguug	ggcucuuccg	agcagcugug	ugcucucggg	gcguggccaa	auccaucgau	2280
uucauccccg	uugagacacu	cgacguuguu	acaaggucuc	ccacuuucag	ugacaacagc	2340
acgccaccgg	cugugececa	gaccuaucag	gucggguacu	ugcaugcucc	aacuggcagu	2400
ggaaagagca	ccaagguccc	ugucgcguau	gccgcccagg	gguacaaagu	acuagugcuu	2460
aaccccucgg	uagcugccac	ccugggguuu	ggggcguacc	uauccaaggc	acauggcauc	2520
aaucccaaca	uuaggacugg	agucaggacc	gugaugaccg	gggaggccau	cacguacucc	2580
acauauggca	aauuucucgc	cgaugggggc	ugcgcuagcg	gegeeuauga	caucaucaua	2640
ugcgaugaau	gccacgcugu	ggaugeuace	uccauucucg	gcaucggaac	gguccuugau	2700
caagcagaga	cagccggggu	cagacuaacu	gugcuggcua	eggecacace	ccccggguca	2760
gugacaaccc	cccaucccga	uauagaagag	guaggccucg	ggcgggaggg	ugagaucccc	2820
uucuauggga	gggcgauucc	ccuauccugc	aucaagggag	ggagacaccu	gauuuucugc	2880
cacucaaaga	aaaaguguga	cgagcucgcg	geggeeeuue	ggggcauggg	cuugaaugcc	2940
guggcauacu	auagaggguu	ggacgucucc	auaauaccag	cucagggaga	uguggugguc	3000

guegecaceg	acgcccucau	gacgggguac	acuggagacu	uugacuccgu	gaucgacugc	3060
aauguagcgg	ucacccaage	ugucgacuuc	agccuggacc	ccaccuucac	uauaaccaca	3120
cagacugucc	cacaagacgc	ugucucacgc	agucagegee	gcgggcgcac	agguagagga	3180
agacagggca	cuuauaggua	uguuuccacu	ggugaacgag	ccucaggaau	guuugacagu	3240
guagugcuuu	gugagugcua	cgacgcaggg	gcugcguggu	acgaucucac	accageggag	3300
accaccguca	ggcuuagagc	guauuucaac	acgeceggee	uacccgugug	ucaagaccau	3360
cuugaauuuu	gggaggcagu	uuucaccggc	cucacacaca	uagacgccca	cuuccucucc	3420
caaacaaagc	aagcggggga	gaacuucgcg	uaccuaguag	ccuaccaage	uacggugugc	3480
gccagagcca	aggececuce	cccguccugg	gacgccaugu	ggaagugccu	ggcccgacuc	3540
aagccuacgc	uugcgggccc	cacaccucuc	cuguaccguu	ugggcccuau	uaccaaugag	3600
gucacccuca	cacacccugg	gacgaaguac	aucgccacau	gcaugcaagc	ugaccuugag	3660
gucaugacca	gcacgugggu	ccuagcugga	ggaguccugg	cagccgucgc	cgcauauugc	3720
cuggcgacug	gaugeguuue	caucaucggc	cgcuugcacg	ucaaccagcg	agucgucguu	3780
gegeeggaua	aggagguccu	guaugaggcu	uuugaugaga	uggaggaaug	cgccucuagg	3840
gcggcucuca	ucgaagaggg	gcagcggaua	gccgagaugu	ugaaguccaa	gauccaaggc	3900
uugcugcagc	aggccucuaa	geaggeeeag	gacauacaac	ccgcuaugca	ggcuucaugg	3960
cccaaagugg	aacaauuuug	ggccagacac	auguggaacu	ucauuagcgg	cauccaauac	4020
cucgcaggau	ugucaacacu	gccagggaac	cccgcggugg	cuuccaugau	ggcauucagu	4080
gccgcccuca	ccaguccguu	gucgaccagu	accaccaucc	uucucaacau	caugggaggc	4140
ugguuagcgu	cccagaucgc	accacccgcg	ggggccaccg	gcuuugucgu	caguggccug	4200
gugggggcug	ccgugggcag	cauaggccug	gguaaggugc	ugguggacau	ccuggcagga	4260
uauggugcgg	gcauuucggg	ggcccucguc	gcauucaaga	ucaugucugg	cgagaageee	4320
ucuauggaag	augucaucaa	ucuacugccu	gggauccugu	cuccgggage	ccugguggug	4380
ggggucaucu	gcgcggccau	ucugegeege	cacgugggac	cgggggaggg	cgcgguccaa	4440
uggaugaaca	ggcuuauugc	cuuugcuucc	agaggaaacc	acgucgcccc	uacucacuac	4500
gugacggagu	cggaugcguc	gcagcgugug	acccaacuac	uuggcucucu	uacuauaacc	4560
agccuacuca	gaagacucca	caauuggaua	acugaggacu	gccccauccc	augcuccgga	4620
uccuggcucc	gcgacgugug	ggacuggguu	ugcaccaucu	ugacagacuu	caaaaauugg	4680
cugaccucua	aauuguuccc	caagcugccc	ggccuccccu	ucaucucuug	ucaaaagggg	4740
uacaagggug	ugugggccgg	cacuggcauc	augaccacge	geugeecuug	cggcgccaac	4800
aucucuggca	auguccgccu	gggcucuaug	aggaucacag	ggccuaaaac	cugcaugaac	4860
accuggcagg	ggaccuuucc	uaucaauugc	uacacggagg	gccagugcgc	gccgaaaccc	4920

cccacgaacu	acaagaccgc	caucuggagg	guggcggccu	cggaguacgc	ggaggugacg	4980
cagcaugggu	cguacuccua	uguaacagga	cugaccacug	acaaucugaa	aauuccuugc	5040
caacuaccuu	cuccagaguu	uuucuccugg	guggacggug	ugcagaucca	uagguuugca	5100
cccacaccaa	agccguuuuu	ccgggaugag	gucucguucu	gcguugggcu	uaauuccuau	5160
gcugucgggu	cccagcuucc	cugugaaccu	gageeegaeg	cagacguauu	gagguccaug	5220
cuaacagauc	cgccccacau	cacggcggag	acugeggege	ggcgcuuggc	acggggauca	5280
ccuccaucug	aggegageue	cucagugage	cagcuaucag	caccgucgcu	gegggeeace	5340
ugcaccaccc	acagcaacac	cuaugacgug	gacauggucg	augccaaccu	gcucauggag	5400
ggcggugugg	cucagacaga	gccugagucc	agggugcccg	uucuggacuu	ucucgagcca	5460
auggeegagg	aagagagcga	ccuugagccc	ucaauaccau	cggagugcau	geueeecagg	5520
agcggguuuc	cacgggccuu	accggcuugg	gcacggccug	acuacaaccc	geegeuegug	5580
gaaucgugga	ggaggccaga	uuaccaaccg	cccaccguug	cugguugugc	ucuccccccc	5640
cccaagaagg	ccccgacgcc	ucccccaagg	agaegeegga	cagugggucu	gagegagage	5700
accauaucag	aagcccucca	gcaacuggcc	aucaagaccu	uuggccagcc	ccccucgagc	5760
ggugaugcag	gcucguccac	gggggcgggc	geegeegaau	ccggcggucc	gacguccccu	5820
ggugagccgg	eccccucaga	gacagguucc	gecuccucua	ugacacacau	cgagggggag	5880
ccuggagauc	cggaccugga	gucugaucag	guagageuue	aaccuccccc	ccaggggggg	5940
gggguagcuc	ccgguucggg	cucggggucu	uggucuacuu	gcuccgagga	ggacgauacc	6000
accgugugcu	gcuccauguc	auacuccugg	accggggcuc	uaauaacucc	cuguagecee	6060
gaagaggaaa	aguugccaau	caacccuuug	aguaacucgc	uguugcgaua	ccauaacaag	6120
guguacugua	caacaucaaa	gagegeeuca	cagagggcua	aaaagguaac	uuuugacagg	6180
acgcaagugc	ucgacgccca	uuaugacuca	gucuuaaagg	acaucaagcu	ageggeuuce	6240
aaggucagcg	caaggcuccu	caccuuggag	gaggcgugcc	aguugacucc	accccauucu	6300
gcaagaucca	aguauggauu	cggggccaag	gagguccgca	gcuuguccgg	gagggccguu	6360
aaccacauca	aguccgugug	gaaggaccuc	cuggaagacc	cacaaacacc	aauucccaca	6420
accaucaugg	ccaaaaauga	gguguucugc	guggaccccg	ccaagggggg	uaagaaacca	6480
gcucgccuca	ucguuuacec	ugaceuegge	gucegggucu	gcgagaaaau	ggcccucuau	6540
gacauuacac	aaaagcuucc	ucaggcggua	augggagcuu	ccuauggcuu	ccaguacucc	6600
ccugcccaac	ggguggagua	ucucuugaaa	gcaugggcgg	aaaagaagga	ccccaugggu	6660
uuuucguaug	auacccgaug	cuucgacuca	accgucacug	agagagacau	caggaccgag	6720
gaguccauau	accaggccug	cucecugece	gaggaggccc	gcacugccau	acacucgcug	6780
acugagagac	uuuacguagg	agggcccaug	uucaacagca	agggucaaac	cugegguuae	6840

agacguugcc gcgccagcgg ggugcuaacc acuagcaugg guaacaccau cacaugcuau 690,0 gugaaagece uageggeeug caaggeugeg gggauaguug egeceacaau cucagaaage 6960 caggggacug aggaggacga gcggaaccug agagccuuca cggaggccau gaccagguac 7020 ucugececue cuggugauce ecceagaceg gaauaugace uggagcuaau aacauccugu 7080 uccucaaaug ugucuguggc guugggcccg cggggccgcc gcagauacua ccugaccaga 7140 gacccaacca cuccacucge eegggeugee ugggaaacag uuagacacuc eecuaucaau 7200 ucauggeugg gaaacaucau ceaguaugeu ecaaceauau ggguuegeau ggueeuaaug 7260 acacacuucu ucuccauucu caugguccaa gacacccugg accagaaccu caacuuugag 7320 auguauggau caguauacuc cgugaauccu uuggaccuuc cagccauaau ugagagguua 7380 caegggeuug aegeeuuuue uaugeacaca uacucucace aegaacugae geggguggeu 7440 ucageceuca gaaaacuugg ggegeeacee cucagggugu ggaagagueg ggeuegegea 7500 gucagggcgu cccucaucuc ccguggaggg aaagcggccg uuugcggccg auaucucuuc 7560 aauugggegg ugaagaccaa gcucaaacue acuccauuge eggaggegeg ccuaeuggac 7620 uuauccaguu gguucaccgu cggcgccggc gggggcgaca uuuuucacag cgugucgcgc 7680 geeegaeece geneaunaen enneggeena eneenaennn negnagggn aggeenenne 7740 cuacuceceg cuegguagag eggeacaeae uagguaeaeu ecauageuaa euguueeuuu 7800 ասուսասան աստասասան աստասասան աստասասան շատաանան աստաշշշաշ 7860 uuucuucccu ucucaucuua uucuacuuuc uuucuuggug gcuccaucuu agcccuaguc 7920 acggcuagcu gugaaagguc cgugagccgc augacugcag agagugccgu aacuggucuc 7980 7994 ucuqcaqauc augu

<210> 9

<211> 340

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic RNA

<400> 9

accugecceu aauaggggeg acaeucegee augaaucaeu ecceugugag gaacuaeugu 60

cuucacgeag aaagegeua gecauggegu uaguaugagu gueguaeage cuccaggeec 120
ceeccucceg ggagageeau aguggueuge ggaaceggug aguacaeegg aauugeeggg 180
aagacugggu ecuuucuugg auaaaeeeae ucuaugeeeg gecauuuggg egugeeeeg 240
caagacugeu ageegaguag eguuuggguug egaaaggeeu ugugguaeug ecugauaggg 300
egeuugegag ugeeeeggga ggueueguag aeegugeaee 340

<210> 10

<211> 340

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic RNA

<400> 10

accegecceu aauaggggeg acaeucegee augaaucacu ceceuguga gaacuacugu 60
cuucacgcag aaagcgucua gccauggcgu uaguaugagu gucguacage cuccaggece 120
cccccucceg ggagagccau aguggucuge ggaaceggug aguacacegg aauugceggg 180
aagacugggu ccuuucuugg auaaacccae ucuaugcceg gecauuuggg cgugceceeg 240
caagacugcu ageegaguag cguuggguug cgaaaggecu ugugguacug ccugauaggg 300
ugcuugcgag ugcceeggga ggucucguag accgugcace 340

<210> 11

<211> 236

<212> RNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic RNA

<400> 11 ageggeaeae acuagguaea cuccauageu aacuguueeu uuuuuuuuu uuuuuuuu 60 uauucuacuu ucuuucuugg uggcuccauc uuagcccuag ucacggcuag cugugaaagg 180 ucegugagee geaugacuge agagagugee guaacuggue ucucugeaga ucaugu 236 <210> 12 <211> 232 <212> RNA <213> Artificial Sequence <220> <223> Description of Artificial Sequence: synthetic RNA <400> 12 ageggeacae auuageuaca euceauageu aacuguuceu uuuuuuuuuu uuuuuuuuu 60 uuuuuuuuu kuuuuucuu uuuuuuuuu uuucccucuu ucuucccuuc ucaucuuauu 120 cuacuuucuu ucuuggugge uccaucuuag cccuggucac ggcuagcugu gaaagguccg 180 ugagcegeau gacugeagag agugeeguaa euggueueue ugeagaueau gu 232 <210> 13 <211> 17 <212> DNA <213> Artificial Sequence <220>

cgggagagcc atagtgg

<400> 13

17

<223> Description of Artificial Sequence: synthetic DNA

```
<210> 14
<211> 19
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 14
agtaccacaa ggcctttcg
                                                            19
<210> 15
<211> 21
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 15
                                                             21
ctgcggaacc ggtgagtaca c
<210> 16
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
```

```
<400> 16
                                                            20
aacaagatgg attgcacgca
<210> 17
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 17
cgtcaagaag gcgatagaag
                                                            20
<210> 18
<211> 30
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 18
gcactetetg cagteatgeg geteaeggae
                                                              30
<210> 19
<211> 28
```

<212> DNA

<213>	Artificial Sequence	
<220>		
<223>	Description of Artificial Sequence: synthetic DNA	
<400>	19	
		28
cccccg	gtgag gaactactgt cttcacgc	20
<210>	20	
<211>	24	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Description of Artificial Sequence: synthetic DNA	
<400>	20	
ccggga	agage catagtggte tgeg	24
<210>	21	
<211>	30	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Description of Artificial Sequence: synthetic DNA	
<400>	21	
ccacto	caaag aaaaagtgtg acgagetege	30

ccactcaaag aaaaagtgtg acgagetege

```
<210> 22
<211> 18
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 22
ggcttgggca cggcctga
                                                            18
<210> 23
<211> 30
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 23
                                                              30
geggtgaaga eeaageteaa aeteaeteea
<210> 24
<211> 21
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
```

```
<400> 24
                                                             21
agaacctgcg tgcaatccat c
<210> 25
<211> 23
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 25
cccgtcatga gggcgtcggt ggc
                                                             23
<210> 26
<211> 27
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 26
                                                              27
accagcaacg gtgggcggtt ggtaatc
<210> 27
<211> 18
```

<212> DNA

```
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 27
                                                            . 18
ggcacgcgac acgctgtg
<210> 28
<211> 30
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 28
                                                                30
agetageegt gactaggget aagatggage
[Sequence Listing Free Text]
SEO ID NOS: 1, 2, 7 and 8 set forth the sequences of replicons.
SEQ ID NOS: 9 to 12 set forth the sequences of synthetic RNAs.
SEQ ID NOS: 13 to 28 set forth the sequences of synthetic DNAs.
[Brief Description of Drawings]
       [Fig. 1]
Fig. 1 is a schematic view showing procedures for constructing a template DNA
```

for preparing the HCV-RNA replicon according to the present invention. The upper section of Fig. 1 shows the structure of the region within pJFH1 or pJCH1, with the viral genome inserted into it. The lower section of Fig. 1 shows the structure of the region within pJFH1 or pJCH1,

with the viral genome inserted into it, that had been constructed by substituting a part of viral genome-inserted region of pJFH1 or pJCH1 with a DNA fragment containing a neomycin resistance gene and EMCV IRES. Symbols in Fig. 1 are as described below. T7, T7 RNA promoter; G, dGTP that was inserted upstream of the 5' end of the inserted DNA derived from JFH-1 or JCH-1 and downstream of the 3' end of T7 RNA promoter sequence; 5' NTR, 5' untranslated region; Core, core protein; and 3' NTR, 3' untranslated region. E1 and E2 represent envelope proteins. NS2, NS3, NS4A, NS4B, NS5A and NS5B represent non-structural proteins. Age I, Cla I and Xba I represent cleavage sites of restriction enzymes Age I, Cla I and Xba I, respectively. GDD, the position of amino acid motif GDD corresponding to the active center of NS5B protein; neo, neomycin resistance gene; and EMCV IRES, internal ribosome entry site of encephalomyocarditis virus (EMCV IRES).

[Fig. 2A]

Fig. 2 shows the nucleotide sequence of rSGREP-JFH1.

[Fig. 2B]

Fig. 2 shows the nucleotide sequence of rSGREP-JFH1.

[Fig. 2C]

Fig. 2 shows the nucleotide sequence of rSGREP-JFH1.

[Fig. 2D]

Fig. 2 shows the nucleotide sequence of rSGREP-JFH1.

[Fig. 2E]

Fig. 2 shows the nucleotide sequence of rSGREP-JFH1.

[Fig. 2F]

Fig. 2 shows the nucleotide sequence of rSGREP-JFH1.

[Fig. 3A]

Fig. 3 shows the nucleotide sequence of rSGREP-JCH1.

[Fig. 3B]

Fig. 3 shows the nucleotide sequence of rSGREP-JCH1.

[Fig. 3C]

Fig. 3 shows the nucleotide sequence of rSGREP-JCH1.

[Fig. 3D]

Fig. 3 shows the nucleotide sequence of rSGREP-JCH1.

[Fig. 3E]

Fig. 3 shows the nucleotide sequence of rSGREP-JCH1.

[Fig. 3F]

Fig. 3 shows the nucleotide sequence of rSGREP-JCH1.

[Fig. 4]

Fig. 4 shows photographs showing the colony formation of Huh7 cells to which rSGREP-JFH1, rSGREP-JFH1/GND and rSGREP-JFH1/dGDD was transfected, respectively. The amount of each of three transfected RNAs in the upper section was 100 ng and that of three transfected RNAs in the lower section was 300 ng.

[Fig. 5]

Fig. 5 shows photographs showing colony formation of Huh7 cells to which rSGREP-JFH1 and rSGREP-JCH1 respectively had been transfected when the concentration of G418 was 0.5 mg/ml of the medium. The amount of each of these RNAs transfected was 100 ng.

[Fig. 6]

Fig. 6 shows photographs showing the effect of Mung Bean Nuclease treatment conducted on the colony-forming ability of the transfected cells. The amount of rSGREP-JFH1 RNA transfected was 100 ng for both cases. The concentration of G418 was 1.0 mg/ml in both media.

[Fig. 7]

Fig. 7 shows photographs showing colony formation when total cellular RNA derived from the replicon-replicating cell clone, which had been established by transfection of rSGREP-JFHI, was retransfected to another Huh7 cells. The photograph on the left shows that the formation of 96 colonies was observed as a result, when using the total cellular RNA derived from the replicon-replicating cell

clone No. 6. The photograph on the right shows that the formation of 77 colonies was observed as a result, when using the total cellular RNA derived from the pool clones. In both cases, RNA was retransfected in an amount containing  $1\times10^7$  copies of the replicon RNA.

[Fig. 8]

Fig. 8 shows photographs showing the results of detecting by the Northern blot method using an rSGREP-JFH1-specific probe for the total RNA derived from a cell clone that had been obtained by retransfecting the total cellular RNA (derived from the replicon-replicating cell clone established by transfection of rSGREP-JFH1) into another Huh7 cells. Explanation of the lanes is as follows. 10<sup>8</sup> represents sample prepared by adding 10<sup>8</sup> copies of the replicon RNA synthesized in vitro to total RNA extracted from Huh7 cells. 10<sup>7</sup> represents sample prepared by adding 10<sup>7</sup> copies of the replicon RNA synthesized in vitro to total RNA extracted from Huh7 cells. Huh7, total RNA extracted from untransfected Huh7 cells; pool clone, total RNA extracted from the pool clones; and 1-11, total RNA extracted from each of cell clones Nos. 1 to 11. "Replicon RNA" represents the electrophoresed position of a molecular weight marker indicating the size of rSGREP-JFH1, "28S" represents the same of a ribosomal RNA marker indicating the size of molecular weight of 4.5 kb, and "18S" represents the same of a ribosomal RNA marker indicating the size of molecular weight of 1.9 kb.

[Fig. 9]

Fig. 9 shows photographs showing the presence or the absence of the incorporation of a neomycin resistance gene into the genomic DNA of a host cell in the cell clone to which rSGREP-JFH1- or rSGREP-JCH1-derived replicated replicon RNA was retransfected. Explanation of the lanes in the photograph on the left is as follows. M, DNA molecular weight marker; 1-8, rSGREP-JFH1-derived cell clones Nos. 1 to 8; N, untransfected Huh7 cells; and P, positive control (PCR amplification product of the neomycin resistance gene). Furthermore, explanation of the lanes in the photograph on the right is as follows.

M, DNA molecular weight marker; and 1-6, rSGREP-JCH1-derived cell clones Nos. 1 to 6.

[Fig. 10]

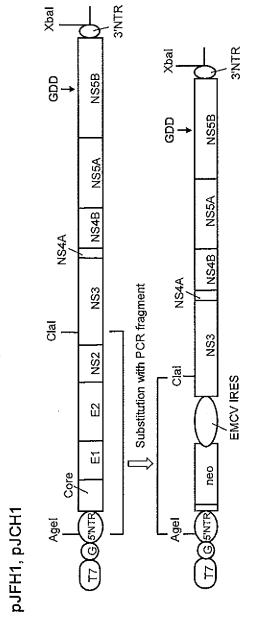
Fig. 10 shows photographs showing the results of detecting NS3 protein expressed in the cell clone that was retransfected with rSGREP-JFH1- or rSGREP-JCH1-derived replicated replicon RNA. Lanes 1 to 8 of the photograph on the left represent rSGREP-JFH1-derived cell clones Nos. 1 to 8. Lanes 1-6 of the photograph on the right represent rSGREP-JCH1-derived cell clones Nos. 1 to 6. Lane P of the photograph on the right represents NS3 protein (positive control) and N represents protein extracted from untransfected Huh7 cells (negative control).

[Fig. 11]

Fig. 11 shows the positions of nucleotide mutations in replicon RNAs obtained from 21 cell clones that were established through the re-transfection of rSGREP-JFH1-derived replicated replicon RNA into Huh7 cells. Mutation positions are indicated using bar lines shown with nucleotide numbers listed in Table 4. A thick bar line denotes nonsynonymous substitution and a thin bar line denotes synonymous substitution.

## [Title of Document] Drawings

[Figure 1]



pSGREP-JFH1, pSGREP-JCH1

# [Figure 2A]

16 Jendaudaro Jendaudaro	50 CCCCUGUGAG	AUGANUCACU	30 ACACUCCOCC	20 AAUAGGGGG	ACCUGECCEU
120 CUCCAGGCCC	J16 GUCGUACAGO	100 Daguavgagu	90 GCCAUGGCGU	08 Anagogoda	70 CUUCACGCAG
180 AAUUGCCGGG	AGUACACOSG	160 GGAACCGGUG	150 AGUGGUCOGO	140 GGAGAGCCAU	130
240	230 GCCAUUUGGG	220 UCUAUGCCCG	01S CADOCAAUA	200 CCUUDOUGG	190 Aagacugggu
OOE OOOKUARKEDO	290 UGUGGUACUG	08S CGAAAGGCCU	i 270 CGURGGUUG	260 AGCOGAGUAG	250 CAAGACUGCU
06E DOAAAUDOWA	350 Augagcacaa	340 ACCGUGCACC	330 GGUCUCGUAG	320 UGCCCCGGGA	310 CGC00GCGAG
420 AUGGAUUGCA	410 AUUGAACAAG	000 DOGCCCAAUG	390 ACACCAACCG	380 ACCAAAAGAA	370 Саладалал
480 CACAACAGAC	470 UAUGACUGGG	460 GCUAUUCGGC	450 GGGUGGAGAG	440 COGCCGCUU	430 CCCACGUUCU
540 cegoucuou	530 Caggggggg	520 GCUGUCAGCE	510 COSUGUUCOG	500 UCUGAUGCOG	490 ANDOGGOGO
CGCGGCUAUC	GACGAGGCAG	QGAAÇQGÇAG	GUGCCCUGAA	550 GACCUGUCCG	UGUCAAGACC
CUGAAGOGGG	GACGUUGUCA	ascuguecuc'	UUCCUUGCGC	069 DODDDOADOK	GUGGCUGGCC
	COCCOGOCAO	OPPOCHOOMO	OCOMMODOCC.	680 CUÇCUAUUGG	WWW
CCCUUGAUCC	Ceccoscava	UGCAAUGCGG	UCAUGGCUGA	760 Anaguaucca	UCCUGCCGAG
					OSCURCOVEC
UCGCGCCAGC	CAUCAGGGGC	<b>SOACGAAGAG</b>	nggaugaucu		028 UDDÓCDRADD
UCBUGACCUA	encontroca :		UNIXULKIKKIN	GCCMGGCGCM	910 CGAACUGUUC
1020 GAUDCAUCGA	1010 CGCUUUUCUG	1000 Sgaaabugge	990 NUNUCAUGGU	980 ASCORUCCOR	970 UGGCGAUGCC
CCCGUGAVAU	Seguuggeua (	icaggacaua i	CEGACOGCUA		CUGUGGCCGG
1140 BUAUCGCCGC	1130 EUGCUUUAÇÇ	coeconocáç (	1110 AAUGGGCUGA	1100 CUNGECGGCG	0e01 Qaqaaqayyyu
1200 PAGUUUNAAC	1190 INGUDOUCU (	1180 CUUCUUGAC	1170 XXVXXXXXXXX	1160 CAGOGCAUGG	1150 UCCCGAUUCG
1260 SGCCOGUGUO	1250 Navaaqquu	1240 CCCGAAGCCG	, 1230 Vacquiyacugi	1220 (CCCCCCCO	1210
				1280 LUUUAOOQUAU	1370 CGUUGUCUA 1
08E1 AADDKAACC	1370 UCCCCUCUC (	1360 Vaggggucu t	1350 NEGAGOARES	1340 IKUCUUCUUG J	1330 AACCUGGCCC I

## [Figure 2B]

				-	
1440	QQAAGCUUCU	1420	gugaagaag	1400	1390
Ugaagacaaa		CAGUUCCUCU	1410	GUUGAAUGUQ	USCAAGGUCU
1500	1490	1480	1470	1460	1450
AGGUGCGACU	ACCUSCOSÁC	GGAACCCCCC	UGCAGGCAGO	AGCGACCCUU	Caacgucugu
1560	1550	1.540	1530		1810
1620 UDCAACAAGG	1610 CUCAAGCGDA	1600 Ducugeucuc	1590 GAAAGAGUCA	. 1580 GAUAGUUGUG	1570 Vogugaguug
1680	1670	1660	1650		រគាត់
1740	1730	. 1720	1710		1690
	1790	1780	1770	1760	1750
	CCAUCACUGC	ACCAUGGCUC	ACACGAUGAU	AAAAQUUUGAAAA	ACGUSGUUU
1860	1850	1840	1830		1810
	1910	1900	1890	1880	1870
	CCUUCCUCGG	GUCUCUCAGU	CCUGUCCACA	Naguccaaan	CAGGCCGGG
1980 C3GCOUACGG	1970 AGACUCUAGC	0961. Adagoondo	1950 UUACCACGGA	1940 UGUGGACUGU	1930 UCGGGGGUUU
2040	2030	2020	cucevenecn	2000	1990
GCCCAGCCCC	December	Gagggggacij		CGCAGAUGUA	GGUCCGGUCA
	2090	2080	2070	2060	2050
	ADAUCOACOU	UGUGGAĞCOĞ	GCCGUGCAAG	AGUCUUXGGA	CCUGGGACCA
	2150	2140	213G	2120	2110
	200000000	CGCGGGACA	GGCUCGGAGA	AUGUCAUCCC	OGGAACGCUG
	2210	2200	011S	2180	2170
	CGGUGGUGUG	000000000	COUDOGRAGO	UUUCGACCUU	COGAGACCCA
2280	2270	2260	225ō	2240	CYCGROGRAGE
NUCCAUCGAU	GCGUGGCCAA	UGCUCUCGGĞ	AGCAGCUGUG	GGCUCUUCCG	3830
2340	DEÉS.	2320	2310	2300	2290
Ugacaacago	DADOUUDADO	Acaaggucuc	CGACGUUGUU	UUGAGACACU	2290000000
2400	2390	2380	2370	.0360	2350
AACÜGGĞAĞU	UGCAUGCUCC	GOCGGGOACU	GACCUAUCAG	CUSUSCOCA	ACCCACCGG
2460	2450	CCCCCCAGG	0848	2420	2410
Diograpauja	Udaaacaadd		Uaudooougu	CCAAGGUÕÕĈ	Ggaragagca
2520	2510	2500	2490	2480	2470
ACAUGGCAUC	UAUCCAAGGC	GGGGGGUACC	CCUGGGGUUU	UAGCUGCCAC	AACCCCUCGG
2580 Cacguacucc	2570 GGGAGGCCAU	03ES. OCCADUADUO	0575 Ngucagacco Ngucagacco	2540 Dekoadeauu	2530 AMUCCCAACA
2640	2630	2620	2610	2600	2590
Caucaucaua	GCGCCUNUGA	UGCGCUAGCG	CGAUGGGGC	AAUUUUUCGC	ACAUAUGGCA
2700	2690	2680	2670	CCACCOCO	2650
GGUCCUUGAU	CAUCGGAAC	POUCUCACO	GGALXCCUACC	USBS	UKCGAUGAAU
2760	2750	2740	2730	2720	

## [Figure 2C]

GUONCAACOX	2780 CCCAUCCCC	2790 Unuagaagag	) 2800 GUAGGCCUCX	) 2810 3 GGCGGGAGGG	2020 UGAGAUCCCC
2830 UUCUAUGGG/	) 2840 A GGGGADUKK	2850 CCUAUCCUGO	2860 AUCAAGGGAG	2870 GGAGACACCI	S880 SUDUUUUUGO
2890 Cacucalage	0002 Aaarakeuga	2910 CGAGCÜCGCG	2920	2930 GGGGCAUGGG	2940 CUUGAAUGCO
GUGGCADACU	3960 UNGAGGGUU	2970 GGACGUCUCC	2980 AUNAUACCAG	2990 CUCAGGGAGA	3000 OUSSUBSUSU
		3030 Gacogogoac			
		3090 UGUCGACUUC			
		3150 UGUCUCACGO			
		3210 UGUUUCCACU			
3250 QUAGUXXUUU	3260 GUSAGUGCUA	3270 CGACGCAGGG	3280 GCUGCGUGGU	3290 ACGAUCUCAC	00ee Dadoodagoa
3310 ACCACCGOCA	3320 GGCUUAGAGC	3330 CUAUUUCAAC	3340 ACGCCCCGGCC	3350 UACCCGUGUG	3360 UCANGACCNU
3370 CUUGAAUDDU	3380 GGGAGGCAGD	3390 UUUCACCGGC	3400 CUCACACACA	3410 WAGAGGCCA	3420 COUCCUCCC
3430 Canacabage	3140 Ameceegega	3450 GAACUUCGCG	3460 UACCUAGUAG	3470 CCUACCAAGC	3480 Unceguesc
3490 GCCAGAGCCA	3500 AGGCCCCCCC	3510 CCCGUCCUGG	3520 GACGCCAUGU	3530 GGAAGUGOÇU	3540 GGCCCGACUC
ANGCOUNCEC	DOGCGGGGCCC	3570 CACACCUCUC	CUGUACCEUO	UGGGCCCUAU	UACCANUGAG
3610 GUCACCCUCA	T820 CACACCCUGG	3630 Gacgaaguac	040E UADADODOVA	3650 GCAUGCAAGC	3660 UGACCUUGAG
3670 GUCAUGACCA	3680 GCACGUGGGU	3690 CCUAGCUGGA	3700 GGAGUCCUGG	3710 CAGCCGUCGC	3720 OGCAUAUUGC
CUGGCGACUG	GAUGCGUUUC	3750 CAUCAUCOGC	CGCUUGCACG	UCAACCAGCG	AGUCGUCGUU
3790 GOGCOGAVA	3800 Aggaggücci)	3810 Guadgaggcu	028E ADAĐUAĐUUU	ocse Doaaddadu Doaaddadu	0440 GOCUCUDODO
		3870 GCAGCGGAUA		3890 Vgaacyccaa	
		0525 • <b>06</b> 222200620		3950 CCGCUAUGÇA	3960 GGCUUCAUGG
9970 CCCAAAGOGG	3980 Aacaauruug	0000 Cocadacoo	1000 UDJADOUGUA	4010 UCAUUAGOGG	4020 CAUCCANUAC
		4050 GCCAGGGAAC (			
4090	4300	4110 GUCGACCAĞU X	4120	4130	4140
* ***********	- W-L-AGAAA	* * ** ** * * * ** * * * * * * * * * *			

## [Figure 2D]

U	•				
0216	4160	4170	4180	4190	4200
UDODAUUDOKU	CCCAGAUCGC	ACCACCCGCG	GGGGCCACCG	GCUCUGUCGU	CAGUGGCCUG
4210	4220	4230	4240	4250	0350
GUGGGGGCUG	4220	CAUAGGCCUG	GGUANGGUGO	UGGUGGAÇAU	ASSACQQUOO
4270	4280	4290	doca	4310	4320
UAUGGUGCGG	GCAUJUCGGG	GCCCCUCGUC	GCAUUCAAGA	UCAUGUCUGG	ÇGAGAAGCCC
0330	4340	4350	4360	4370	4380
DKADQUAUQU	AUGUCAUCAA	UCUACUGCCU	GGGAUCCUGU	CUCCGGGAGC	CCUGGUGGUG
4390	4400	4410	6420	4430	CGCGGUCCAA
GGGDCAUCU	GCGCGGCCAU	UCUGOGCOGC	ÇAÇGUGGGAC	CGGGGGAGGG	
4450	4460	4470	0480	4490	. 4500
Coanganaca	GCUUNUUGC	CUUUGCUUCC	Dokarđara	ACGUCGCCC	UACUCACUAC
Ø510	4520	4530	0445	4550	4560
GLYJACEGAGU	CGGAUGCGUC	GCAGOGUGUG	Daugaaggaa	00GGCDCDCU	UACUAUAACC
4570	4580	0998	0004	4610	4620
AGCCUACUCA	GAAGACUCCA	Auadrugana	UOKOĐAĐUOA	GCCCCAUCCC	AUGCUCCGGA
4630	4640	4650	DCCACCAUCU	4670	4600
VCCVGGCVCC	GCGACGUGUG	GGACUGGGOU		UGACAGACUU	Caaaaauuugg
4690	4700	4710	60000000000000000000000000000000000000	4730	4740
CUCACCUÇUA	ANUQUUCCC	CANGCUGGGG		TRAUCUCUUG	UCAAAAGGGG
4750	USUGGCCGG	4770	4780	4790	CCCCCCCAAC
Dacaagggug	USUGGCCGG	CACUGGCAUC	AUGACCACGC	CUCCCCUIG	
4810 AUCUCUGGCA	AUGUCOGCCU AUGUCOGCCU	GGGCUÇUAUG		GGCCUNAAAC	CUSCAUGAAC
4870 ACCUGGCAGG	4880 CCACCUUCC	UNUCANUUGC	,		
	A940 ACAAGACOGC	•			
	5000 CGUACUOCUA				•
	5060 CUCCAGAGUU		GOGGACGGOG		
	AGCCGUUUUU AGCCGUUUUU		GUCUCGUUCU	-	
	5180 CCCAGCUUCC				•
CUANCAGAUC	CGCCCCACAU		VCDECCCCC	GGCGCUUGGC	ACGGGGAUCA
CCUCCAUCUG	0065 DUDDAGDODA	CUCAGUGAGC	CAGCUAUCAG	CACCGUCGCU	ecedectact
5350	5360	5370	OSEE	5390	5400
UGCACCACCC	ACAGCAACAC	CUADAACGUG	DOUDDUADAD	NGCCAACCU	GCUCAUGOAG
5410 GCCGGUGUCG	OSBE CLACAGACAGA	5430 GCCVGAGUCC	5440 Agggugecog	5450 UUCUGGACUU 1	5460 CCXCGAGCCA
5470	5480	5490	5500	5510	5520
MUGCCGAGG	Angngagga	ČČUUĞAĞÇÇÇ	UCANUACCAU	CGGAGUGCAU	GCUCCCCAGG

## [Figure 2E]

				•	_
5580	5570	5560	5550	5540	5530
GCCCCUCGUG	ACUACAACCC	GCACGGCCUG	ACCCGCCUGG	CACGGGCCUU	AGCGGGUUC
5640	5630	5620	5610	5600	0068
UCUCCCCCC	CUGGUUGUGC	CCCACCGUUG	UUACCAACOS	GGNGGCCAGA	Ganucuuga
5700	. 5690	AGACGCCGGA	5670	5660	5650
Gagogagage	CAGUGGGUCU		UCCCCCANGG	CCCCGACGCC	CCCANGAAGG
	5750	5740	5730	5720	5710
	UUCGCCAGCC	AUCAAGACCU	GCAACUGGCC	AAGCCCUCCA	ACCAUAUCAG
9ACGUCCCCU	5810	5800	5790	5780	OTTE
	CCGCCGGUCC	GCCGCCGANU	GGGGGGGGGC	GCUCGUCCAC	DADĐUAĐUĐA :
5880	5870	5860	5650	5840	5830
CGAGGGGAG	VGCCCCCCCU	GCCUCCUCUA	GACAGGUUCC	CCCCCUCAGA	GGUGAGCGG
5940	5930	5920	5910	5900	5890
CCAGGGGGGG	AACCUCCCCC	GUAGAGCUUC	GUCUGAUCAG	CGGACCUGGA	CCUGGAGAUC
0000	5990	5980	5970	5960	5950
OCAUADOASO	GCUCCGAGGA	UGGUCUACUU	CUCGGGGUCU	CCGGUUCGGG	GGGGUAGCUC
0808	6050	0100	0603	6020	6010
22228AUSUS	Daauaacucc	DUDDDDDDDDDA	DDDDDDDAUA	GCCCAUGUC	ACCGUGUGCU
6120	6110	6100	6090	6080	6070
CCAUAACAAG	DGUUGCGAUA	AGUAACUCGC	CAACCCUUUG	AGUKGCCAAU	Gaagagaaa
6180	6170	6160	6150	6140	6130
UUUUGACAGG	AAAAGGUAAC	CAGAGGGCUA	GAGOGCCUCA	Caacaucaaa	GUSUNCUKUN
•	6230	6220	6210	6300	6190
	ACAUCAAGCU	GKKKULAAAGG	UUAUGACUCA	DOBACGCOCA	ACGCAAGUGC
	•	6280 Gagggggggggggggggggggggggggggggggggggg	6270 CACCUUGGAG	CAAGGCUCCU	6250 AAGGUCAGCG
•		01E3 GAGOCOUSSAS	6330 CGGGGCAAG	6320 AGUAUGGAUU	6310 GCAAGAUCCA
0SP9	6410	6400	09E6	NGCCCCUGUG	6370
ADMOCOCACA	Cacaaacacc	CUGGAAGACC	DVDÓAÐDKAÐ	6380	AACCACAUCA
. 6480 Uaagaaacca	6470 CCAAGGGGGG		GGUGUUCUSC	6440 CCANANNSA	6430 ACCAUCAUGG
6540 GGCCCUCUAU			0510 UGACCUCGGC	0500 UCGUFUACCC	0006 GOUCGCOUCA
6600	6590	0826	6570	6560	6550
CCAGUACUCC	CCUNUGGCUU	DUQQABQQVA	UCAGGOGGUA	AAAAGCUUCC	GACANUACAC
CCCCAUGGGU	addiadaaaa	GCAUGGGGG	6630 AAAQUUQUDU	AUĐAĐĐŲĐĐĐ	CONGOCCNAC
CYGGACCGYG	agagacau	ACCISTICACTIG	6690 Cuucgacuca	AUACCCGAUG	DULUCGUALIG
6780	6770	6760	6750	6740	6730
ACACUCGCUG	GCACUGCCAU	Gaggaggggg	CUCCCOGCCC	BUDDOORDOA	Gaguçcavau
0484	0683	68 <b>20</b>	6810	ÓOSA	6790
Daurdeooxid	DAAACUDDDO	UUCAACAGCA	AGGGCCCAUG	DDAUDOAGGU	ACX8AGAGAC
6900	6890	6880	6870	6860	6850
Cacaugguau	Guaacaccau	DOLLACORODA	COAAUCGHDD	60900AG03G	AGACYUIGCC

# [Figure 2F]

6910 GHGAAAGCCC	DAGCGGCCUG	6930 Caaggougos	6940 GGGAUAGUUG	6950 CCCCACABO	GOUGUAUGO ODUAUGOUOO
6970	6980 11000136130811	6990 CUCAGAAAGC	7000 CNGGSG6CT33	7010 * necesses	7020 GOSGNACCOS
7030 Agagocuuca	7040 CGGAGGOCAU	. 7050 GACCAGGUAC	7060 UCUGCCCCUC	7070 CUGGUGAUCC	7080 CCCCADACCG
2090	7100	7110	7120	7130	7140
GAAUAUGACC	OGGAGCUAAU	AACAUCCUGU	UCCUCAAAUG	UGUCUGUGGC	GUUGGGCCCG
7150 CGGGGCCGCC	7160 GCAGAVACUA	7170 CCUGACCAGA	7180 GACCCAACCA	7190 CUCCACUCIC	7200 COGGGCUGCC
7710	<b>マケン</b> 的	7230	7240	2250	. 2260
UCOGAAACAG	UUAGACACUC	CCCUAUCAAU	UCAUGGCUGG	GAAACAUCAU	CCAGUAUGCU
7270 CCAACCAUAU	7280 GGGGGGCAU	7290 GGUCCUAAUG	7300 ACACACUÚCU	7310 UCUCCAUUCU	7320 CAUGGUÇÇAN
7330		7350			
GACACCCUGG	DOPER ACCAGANA	CAACUUUGAG	AUGUAUGGAU	CAGUAUACUC	CGUGAAUCCU
7390	7400	7410	7420	7430	7440
UUGGACCUUC	CAGCCAUAAU	UGAGAGGUUA	CYCCCCCOINC	ACGCCUUTOUC	UAUGCACACA
7450			7480		
nychcacycc	ACGARCIGAC	CCGGGUGGCU	DCAGCCCUCA	GAAAACIJIXKI	GGCGCCACCC
7510 UZUGGGGAZYZ	7520 GGANGAGUCA	7530 GGCUCGCGCA	7540 GUCAGGGGGU	7550 CCCUCABQUC	03av Sodadsueco
2570	2580	7598	7600	7610	7620
AAAGCGGCCG	DIMINGCOGCCG	AUAUXUCUUC	AAUUSGGCGG	UGAAGACCAA	GCUCAMACUC
7630 ACUCCAUUGC	7640 CGGAGGCGCG	7650 CCUACUGGAC	7660 UUAUCCAGUU	7670 GGUUCACCGO	7680 OGCGCCGGC
7690		7710			
GGGGGGACY	UUUUUCACAG	CGUGUCGCGC	eccoencecc	GCUCAUUACU	COUCOGOCUÑ
		7770 AGGCCUCUUC			
CUCCUACUUU	UCCUAGOGGU	AGGCCUCUTUC	CUACUCCCCG	CUCGGUAGAG	CGGCACACAC
7810	7820	7830	7840	7850	7860
UNGGUACACU	CCAUAGCUAA	CUGUUCCUUU		UUUUUUUUU	COCOMBUNI
7870	7880	7890	7900	7910	7920
	CUUUUUUU	UUUUUCCCUC	unicunceçu	UĆUĆNUĆUUA	UUCUACUUUC
7930	7940	-7950	7960	7970	7980
		AGCCCUAGUC			
000С Элээнэлэйа	0008 UDOODUDAGA	OLO8 DUDUPPUDAK	8020 UCUGCAGAUC	020\$	8040

## [Figure 3A]

				-	-
	50 CCCCUGUGAG	40 AUGNAUCACU	30 ACACOCOGO	20 Dauagegeco	ACCOGCCCCO
120 CUCCAGGCCC	GUCGUAÇÃGC	100 Uaguaugagu	GCCVNGGCGN 90	NAVGCGGGGAV	70 CINCACGCAG
1.80	170	160	150	140	130
AAUUGCCGGG	AGUACACCGG	GGAACCGGUG	AGUSGUCUĞC	GGAGAGCCAU	CCCCCCCCCC
		220 UCUAUGCCCG	210 ADAMACCCAC	200	190 Aagackeescu
			270 CGUUGGUUG	098 Daudadooga	250 CAAGACUGCU
ÓĞE	350	340	330	320	310
DOMADOOUK	Adgagcacaa	ACCGUGCACC	GGUCUCGUAG	UGCCCCGGA	UGCUUGCGAG
	410	400	390	380	370
	AUUGAACAAG	Vogeogaavg	ACACUAACCG	ACCAAAAGAA	UCANAGAAAN
	470 Unusacusse	460 \$\$\$\$\$\$\$\$\$\$\$\$\$\$	450 GGGUGGAGAG	COGGCCGCUU	430 CGCAGGUUCU
540	530	920	510	500	190
CGGUUCUUUU	CXCCCCCCCC	ÇÇVÇVÇAÇÇÇ	CCGUGUUCCG	DOOQUADQOOG	AAUCGGCUGG
	590	580	570	560	daa
	Gaccagecag	Ugaacugcag	AAGUXXXXXXX	GACCUGUCCG	Coadaanugu
	650	040	02999000000	620	610
Cucaacces	GACGUUGUCA	OUOQUOQUO	04999000000	ACGACGGGCG	GUGGCUGGCC
CUCACCOUGC	710	700	690	680	670
720	CUCCUGIKAU	UADDASDDDD	GCBAAGUGCC	Cuscuauuso	Dendagora
780 CGCUUGAUCC	770 CGGCUGCAIJA	760 December 1600 Table	750 UCAUGGCUGA	740 Accountera	730 UCCOGCOGAG
	930 DACCOĄCCAC	ACAUCGCAUC	018 AGCARGOGAA	CCAUUCGACC 600	790 SOUDOAUSE
000 22A22222QU	890 Caucaggggc	orb Dadaacoard Ordaacoard	870 Aggaugaucu	GARZACENDG 860	028 UEGCCDBAADD
030	950	940	930	920	910
OCCUGAÇÕÇA	DXVXVAGAGO	GCCCGAOGGC	Vadoododda	GCCAGGCUCA	CGAACUSTUC
1020	1010	0001	000	08Q	970
GAUUCAUCGA	CGCUUUUCUG	Dəğijaaaaqə	UDDŲKOJAŲK	KDÇDƏVUDƏN	Descenuece
1080	1070	1060	1050	1040	1030
CCOGUGAUAU	GOGUUGGOUA	UCAGGACAUA	CGGACCGCUA	CUXSGUGUGG	CUGUGGCCGG
GUAUCGCCGC	GUGCUUUACG	ccecanceae	1110 AAUGGGCUGA	CUUGGCGGCG	UGCUGNAGAG
OOŞI	2190	1180	1170	1160	1150
Daaauuugad	UXUUXUU	DAÐUUNUDD	CCUAUCUAUCO	CACCGCAUCG	DOCCGAUDOG
1260	1250	1240	1230	1220	1210
GCCCGUGUG	Cikigaavaa	SCCCAAGCCG	AACCUUACUG		CCUCUCCCUC
1320	1310	1300	1290	1280	. 1270
AGGGCCCGGA	UGGCANUGUG	UGCOGUCUUU I	UCCACCAUAU	UAUGUUAUUU	CGUUGUCUA
1380	1370	1360	1350 DUUADAGOA	1340	1330

## [Figure 3B]

· 1394 UGCAAGGUC	001. Juduaarkuud u	1410 Gugaaggaag	1420 CAGUUCCUCU	) 1430 J GGAAGCONCU	1440 OGANGACANA
CANCGUCUG	) 1460 Jagogačocki	( (GCAGGCAGC	: GGNACCCCCC	: Accussosac	i yegneconcin
151( GOGGCCAAN	) 1520 A GCCACGUSUA	1.530 UAAGAUACAC	2540 CUGCANAGGO	1650 GGCACAACCC	1550 Cagugocacu
UUGUGAGUUK	1580 GAŅĀGUUGUS	GAAAGAGUCA	ANUGGCUCUC	CUCAMBOGUA	UUCAACAAGG
	) 1640 LUGCCCAGANG		-		
	1700 AUGUGUUUAG				
akrudombok	1760 CCUUUGAAAA	ACACGAUAAU	, yccyndddcc	CCAUCACOGC	UUACGCCCAG
	1820 GUCUCUUGGG				
	1880 AGGUCCAAGU				
	1946 UNUXGACUGU				
GECCCGGUCA	2000 CGCAGAUGUA	COCCEAGCGCC	GAGGGGGACU	UGGUCGGGUG	GCCCAGCCCU
CCARROYCCY	DOOK ANUCUNUKSOA	GÇÇGUGUAÇG	UGUGGAGCGG	TICGACCUSUA	INUGGUCAÇO
	2120 AUGUCAUCCC 2180				
CCGAGACCCC	UUUCGACCOU	GAAGGGGUCC	,UCGGGGGGAC	CUGUGCUUUG	CCCUAGGGGC
CACGCOGOCO	2240 GAAUCUUCOG	OGCHACHOOG	THICKNOONS	COCOMSCOM	SUCCAGASAG
DICAUCCCC	2300 UXXXAGACGCU	CCACAUCGUC	ACGCCGUXXXC	CCACCUUUAG	DGACÁACAGO
ACACCACCAG	2360 CUGUGCCCCA	GACCUAUCAG	GUGGGGUACU	UGCACGCCCC	CACUGGCAGU
-	2420 CCAAGGUÇČĆ				-
AAUCCCUCGG	2480 UGGCUGCCAC	CCUGGGAUUU,	GGGGGGACA	UGUCCÁAGGC	acauggcauc
AACCCCAACA	2540 UUAGGACUGG	AGUCAGAACU	GUGACGACCG	GGGNGCCCAU	DACADACUCC
0025 Augguauggaa	2600 NAUUCCUOGC	2610 CGAUGGGGGC	2620 UGCGCX6GCG	2630 GOGCCUAUGA	2640 CAUCAUCAUA
0665 Uaaduaddu	2660 GCCACUCUGU	2670 Geauscuacc	2680 ACUAUUCUOG	2690 GCAUCGGGAC	0072 Dabindouda
. 2710 CAAGCAGAGA	2720 Cagcegegu	DEFE UDAAUDOOAD	2740 GUACUGGCCA	2750 CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	2760 CCCCCCCCCCCC

# [Figure 3C]

. •	•				
EDGACARCO	0 2760 C COCADCCCAA	091S Daddadauau	CONSCIONATION :	Gacagoageg	2820 UGAGAUCCOC
VUCUAUGGG	D 2040	2850	obbs	2870	2880
	A GGGCGUUVCC	CCUGUCUUAC	Dadddaachuk	GGAGGCACUU	GAUUUUĆUGO
CACUCAAG	0 2900	2910	OSES	2930	2940
	A AAAAGUGUGA	CGAGCUCGCA	DUDDODDOA	GGGGCAUGGG	Cuugaacscu
2950	o 2960	2970	0892	299Ó	3000
GUGGCAUAU	O ACAGAGGGUU	GGACGUÇUCC	AACCAUAAUA	COCAAGGAGA	UGUGGUGGUC
3010	osoe	3030	0400	3050	3060
GUUGCCACCX	Dationogoa e	Gacgggggay	UOAQAGQUOA	UCCACUCOCU	OAUCOACIKIC
3070	0806	0000	1190	3110	3120
AACQUAGOGA	088ASCCASSO 7	OUUOADAUDO	AGCCUGGACC	CCACCUCAC	UAUAACCACA
3130	) 3140	3150	3160	3170	3180
CAGACUGUCO	COCAAGACGC	UGUÇUÇÂGGÜ	AGUCAGOGOC	GAGGGGGGAC	GGGUNGAGGA
3194	0026	UGOOUCCACO	OSSE	3230	3340
AGACOGGCI	AUDDAUNUU A		GGDGAGCGAG	CCCCAGGAAU	GUUUGACAGU
3250	) 3260	3270	3280	3290	3300
GUAGUACUCI	J GUGAGUGCUA	CGACGCAGGA	GCUGCUUGGU	NUGAGCUCUC	ACCAGUGGAG
3310	) 3320	3330	3340	3350	3360
ACGACCGUCA	A GGCUCAGGGC	GUAUUUCAAC	AOGCCUGGCU	Decodeusus	CCAGGACCAC
3370	00886	3390	3400	3410	3420
CUUGAGUUD	UDAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	UUUCACCGGC	CUCACACACA	ACKIDƏDADAU	1000000000
3430	) 3440	3450	3460	3670	3480.
Cagacamago	Cagucogogga	AAAUUUOSCA	Unccunguag	CCUAUCAGGC	CACAGUGUGC
3490	9500	3510	3520	3530	3540
GCCNGGGCCI	AAGCGCCCCC	000000000	Gacchcaucu	Genaguecou	Gacucuacuc
3550	) 3560	3570	COGUACCOUÚ	3590	3600
AAGCCCACGO	C. UUSUGĢGCCC	UACACCUCUC		UGGGCUCUGU	Daccaacgag
) i de	J620	3630	048E,	3650	3660
Kudooacked	CACACCCOGU	Gacaaaauac	UKOKOOQOUK	GCAVGCAAGC	DGACCUCGAGU
3674	) 3890	3690	3700	3710	OGCGUAUUGC
GUCAUGACCI	A GCACGUGGGU	CCUGGCUGGG	Geagueuuag	Cagccguogc	CGCGUAUUGC
3730	3740	3750	3760	3770	3780
UUAGQGACQ0	GGUGUGUJUC	CAUCAUUGGC	CGUUUACACA	UCANCCAGOG	AGCUGUCĜUĜ
3790	0088	3910	3820	őebe	3840
3790	UJOUDDARDA /	CUAUGAĞĞCÜ	UUUGAUGAGA	Dukaddaddu	UGCCUCCAGA
ecaecacaca	: UUGAAGAGGG	GCAGCGGAUA		ugaaguccaa.	
3910	3920	0191	3940	950	3960
OUNUUGCAGO	AAGCCUCUAA	ACAGGCCCAG	Gacauacaag	ACGUGUCCO	AGCUUCGUGG
3970	3980	000C	000}	4010	4020
CCCAAGAUGG	AGCAAUUCUG	UKDAAADDQ	Waaqoudua	UCAUAAGOGG	CAUUCAGUAC
4030	4040	4030 GCCAGGGAAC	4060	4070	4080
4090	4100	4110	4120	4130	6140
GCCGCCCCC	CCAGUCCGUO	GUCAACUAGC	ACCACCAUCC	OUCUNARCAU	UCUGRIGGEC

## [Figure 3D]

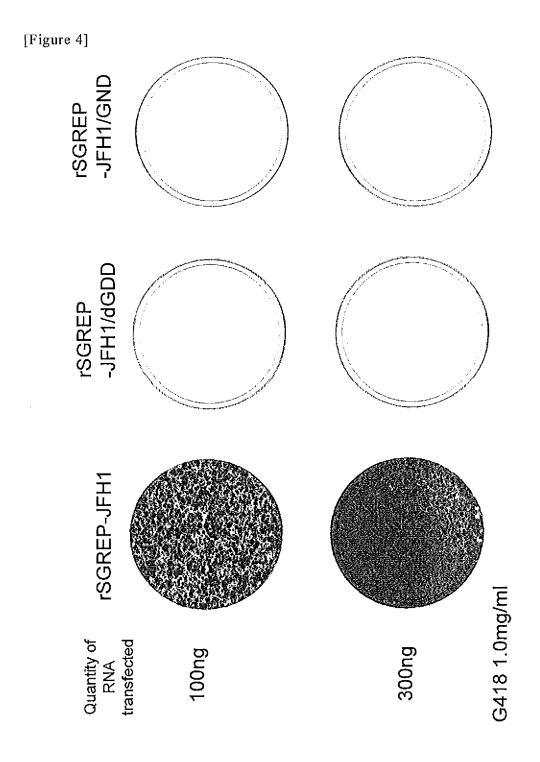
6150	4150	4170	4180	6CONUGUUGU	4200
UGGCUCGCGC	CCCAAAUUGC	GCCACCCGCG	GGGGCCACUG		CAGUGGCCUG
4210 GUCGAGCUG	4220 CUGUUGGCAG	4230 CAUAGGCUUG	4240 GGUAAAGUGC	4250 UGGUGGACAU	4260 CCUGGCAGGG
4270	4280	4290	4300	4310	4320
UAUGGUSUGG	GCAUUUCGGG	66000U00U0	GOGUUUAAGA	UCAUGUCUSS	CGAGAAGCCC
4330	OPED	4350	4360	4370	4380
UÇÇA(KĞAĞĞ	Kananena	COUGCUGCCU	GGGAUUCUGU	CUCCAGGUGG	UCUGGUGGUG
4390	4400	4410	4420	4430	4440
GGAGUCAUCU	GO2CGGCCAU	0003030030	Cangugggac	CGGGGGAAGG	CGCGGUCCAA
4450	4460	4470	4480	4490	4500
UKGAUGAACA	GGCUUAUCGC	cuticgcouce	Agaggaaacc	ACGUCGCCCC	UACUCACUAC
4510	4520	4530	4540	4550	4560
GUGACGGAGU	COGNUCCOUC	GCAGOGUGUC	ACCCAACUGC	UUGGCUCUCU	Cacuauaacu
ás70	4580	0590	4600	4610	4620
Agexuacuca	Geagacouca	Caacusgauc	ACUGAGGAUU	GCCCCAUCXC	AUGOGCCGGC
0630	4640	.4690	4660	073 <b>)</b>	4580
OCCUCCOCOCOCO	GCGAUGUGUG	GGACUGGGUC	Uguaqqavçç	UUDAQAQAAU	Uangaacuss
4690	4700	4710	4720	4730	4740
CUGACCUCCA	AGCIKBUKCC	AÁAGAUGCCU	GGCCUCCCCU	000000003	CCAAAAGGG
4750	4760	4770	4780	4790	4800
VACANGGGOG	UGUGGGCCGG	CACUGGCAUC	Augaccacac	GAUGCCCCIG	0990900AAC
4810	ACCUCCCCUU	4830	0840	485G	4860
AUCUCUGGCA	4820	GGGCUCUAIXG	Daocouarda	Gacciaaaac	CUGCAUGAAC
4870	4880	4890	4900	4910	GCCGAAACCC
ACCUSSCAGS	GGACCUUUCC	UAUCAAUOGU	UAUACAGAAG	0000000000	
4930	4940	4950	4960	4970	rred
Goguuaacu	UCAAGACOSC	CAUCUGGAGA	GUGGCGGCCU	CAGAGDACGO	Dokoudaado
4990	5000	5010.	5020	. 5030	SO40
CAGCACGGAU	Cauaugcoua	UAUAACAGGG	CHEACCACAG	ACAACUUAAA	SOUCCOUCC
5050	5060	5070	obog <sub>ą</sub> cycac	5090	5100
CANCUCCCU	COCCAGAGUU	UUUCUCUUGG	Geogącycac	UACAAAUCCA	UAGGUCCGCC
01.18	5120	5130	. 5140	5150	5160
AADDADADDD	AGCCGUUUUU	CCGCGAUGAG	GUCUCGUUCA	GOGUUGGGCU	CAAUUCAURU
5170	5180	5190	5200	5210	5220
GUCGUCGGG0	CUCAGCUUCC	CUGUGÁCCCO	Gagocogaca	CUGAGGUAGU	Gauguēčānūg
5230	5240	5250	5260	5270	5280
CUNACAGACC	CAUCCCAUAU	CACGGGGGAG	GCUGCAGCGC	ODAUUUBOOD	GCGGGGGUCA
5290	5300	5310	5320	5330	5340
CCCCAUCUG	AggCaagCUC	CUCAĞOĞAĞC	CAGCUGUCGO	CGCCAUCGCU	GCGAGCCACC
5350	5360	5370	5380	5390	5400
UGCACCACCC	ACGGUAGGAC	CUAUGAUGUG	Gacauggugg	AUGCCAACCU	Guycaugggg
5410		5430	5440	5450	5460
5470	5480 AAGAGGGGA	5490	5500	\$510	5520
AUGACOGAGG	NO POSTORONA	ぐていいいいいついかい	OFUSPICATION		<b>シャルル・イン・ファン</b>

## [Figure 3E]

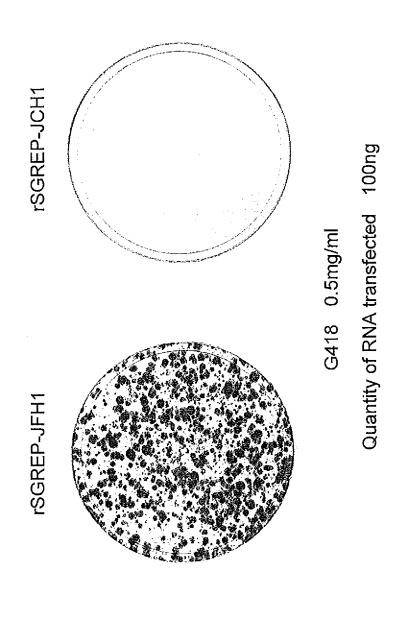
					5580 SUDVINCEOOQA
					5640 UCUCCCCCCC
5650 CCCAAAAAGA	5660 CCCCSACCCC	5670 UCCUCCANGG	5680 Ağacgcogga	5590 Cagugggucu	5700 GAGCGAGAGC
					5760 CCCCCAAGC
5770 GGCGAUUCAG	5760 GCCUUUCCAC	5790 GGGGGGGAC	\$800 GCCGCCGACU	5810 COGGCGAUGG	5820 GACAÇOCCCU
5830 GACGAGUNGG	5840 CUCUUUCGGA	5850 Gacagguuçu	5860 ACCUCCUCCA	9870 UGCCCCCCU	OSAGOGGAGO
5890 CCUGGGGACC	5900 Cagaccuggā	5910 GCCUGAGCAG	0005 Guagagouc	5930 AACCUCCUCO	0466 0000000000
5950 Gaggeageur	5960 COGGCUCGGA	5970 CUCGGGGGCC	0896 UVDAVÇUÇEKI	5990 GCUCCGAGGA	GGADGACOCC
6010 6010	6020 GCUCCAUGUC	6030 AUAUUCCUGG	5040 ACCOGGGCUC	0e06 Doudaauaau	0600 DOCODAUDUU
6070 Салсасска	080a UAADOSUUSA	0990 5090 (AAACUKKUUS	5100 AGCAACUCGC	6110 UGUUGOGAÇA	6120 CCAUAACAAG
6130 GUAUACUGUA	CUACAUCAAA	6150 Gagugocuca	6160 CUANGGGCUA	6170 Aaaagguaac	0180 SQAUASXXVUU
		01S5 ADOUAQUADU			
		6270 CACCUUAGAG			
		0539 Degogodaag			
					6420 ARUUCCUACA
ACCAUCAUGG	6440 CCAAAAAUGA	6450 GGUGUUCUGC	6460 GUGGACCCCG	6470 CCAAGGGGGG	6480 ADDAGAAAUU
		6510 UGACCUCOGC			
		6570 UCAGGGGGGG			
0.699 CCCGCUCAGC	GGGUGGAGUU		GCNUGGGGGG	aaaagaga	CCCUNUGGGU 6660
	AUACYCGAUG	0699 CULUÇAÇUCA	<b>ACCOUCACUS</b>	agagagacau	CAGGACUGAG
6730 Gaguccauau	0176 BUDDDDADOA	GUCCUUACCC	6760 Caggagaca Caggagaga	6770 GAACUGCCAU	6780 ACACUXGCUG
6790 ACUQAGAĞAC	6800 Octangueses	0183 Düadoodeeaa	6820 ACDRACAGON	0£83 OUÐACOÐQQA	6840 CUGCGGGUAC
6850 AGGCGOUGCC	0383 0000K30303	6870 COACUDACC	6880 ACUAGUAUGO	óésá Uaž:dadaadg	6900 CACAUGCUAU

# [Figure 3F]

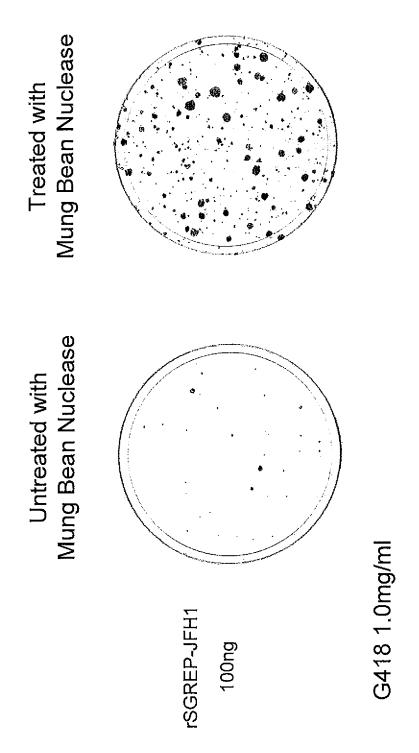
GUAAAAGCCC	) 6920 UAGCCCCUUC	6930 CANGGCUGC	) des Kulaauagoo e	) 6950 } CGCCCACGAT	6960 Figurgauge
6976	6986	6996	7000	7010	7020
GCCOACGACL	UGGUCGUCAU	CUCAGAAAGO	CAGGGGAÇU	AGGAGGACGA	GCGGAACCUG
. 7030 AGAGCCUUCA	7040 CGGAGGCUAU	7050 GACCAGGUAU	7060 UCUGCCCU	7070 CUSGUGACCC	7080 CCCCAGACCG
OPOV COADVAUKAD	7100 UGGAGCDAAU	AACAUCOOGU AACAUCOOGU	7120 UCCOCAAACG	7130 UGUCUGUGG	ACUUGGCCA
7150	7160	7170	7180	7190	7200
CVCCCCCCCCC	GCAGADACUA	CCUGACCAGA	GACCCCACCA	CULCABULEC	CCGGGCUGCC
7210	7220	7230	7240	7250	7260
UGGGAAACAG	UUNGACACUC	CCCUGUCAAU	UCAUGGCUGG	GAAACAUCAU	CCAGUACGCU
7270	7280	7290	7300	7310	7320
CCAACCAUAU	GOGUUCGCAU	CCUCCUSION	ACACACUUCU	DCGCCNDDCG	CAUGGCCCAG
7330	7346	2350	7360	7370	7380
GYCYCCCAYG	ACCAGAACCU	UAACUUUGAA	AUGUACGGAU	COGUGUÁCUC	CGGSYCCGG
7390	7400	7410	7420	7430	7440
	CAGCCAUAAU	UGAAAGGUDA	CACCOCCUUG	ACCCCTOCOC	UKUGCACACA
7450	7460	7470	7480	7490	7500
UACACUCCCC	ACGAACUGAC	eceesusecu	UCAGCCCUCĂ	GAAAACUUGG	GGCGCCACCC
7510					7560
				CCCUCAUCUC	
7570	7580	7590	7600	7610	7620
				UGAAGACCAA	
7630	7640	7650	7660	7670	7680
				GGUUUAÇÕGŬ	
7690	7700	7710	7720	7730	7740
				GCCUAUUACU	
7750	7760	7770	7780	7790	7900
CUCCUACUUU	CUGUAGGGGU	AGGCCCCCCCC	CUACUCCCCG	CUCGAUAGAG	CGCACACÃO
7810	7820	7830	7840	7850	2860
UAGCUNCĂÇŮ	CCAUAGCUAA	CUGUUCCUUU	บบบบบบบบบ	785 <b>0</b> UUUUUUUUU	บบบบบบบบบัง
7870	7880	7890	2000	7910	7920
UTULERARIO	COOCUUOOO	UUUÚUCCCÚC	υυνουσόσου	UCUCAUCION	DUCUACUOUC
3930	7940	7950	7060	<b>ን</b> ዓንስ	7980
UUUCUUGGUG	GCUCCAUCUU	agoccuaguc	ACGGCUNGCO	GUGAAAGGUC	CGUGAGCČĞČ
i dan	ቋስስስ	8010	กรกร	8030	9000
AUGACUSCAG	AGAGUGCOGU	<b>AACUGGUCUC</b>	UCUGCAGAUC	DUGU	- 4444



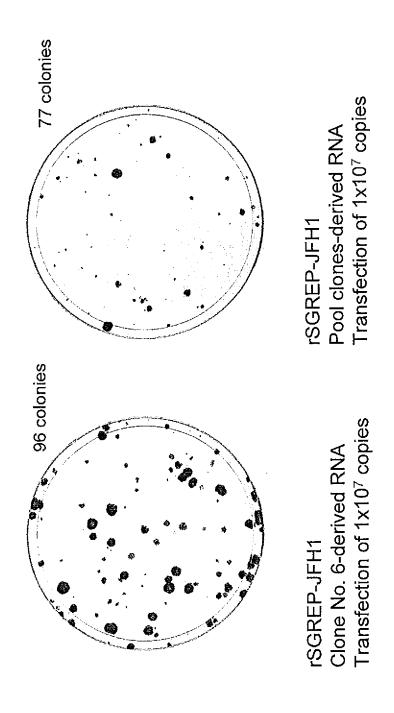
[Figure 5]



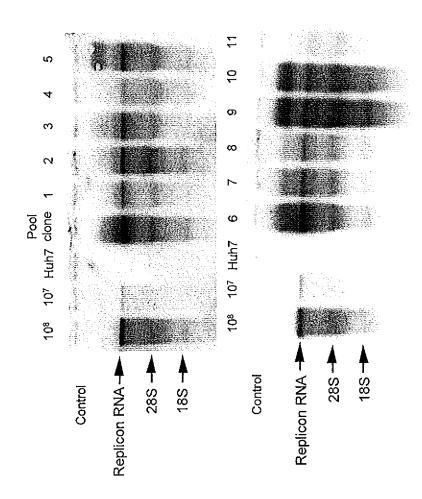
[Figure 6]



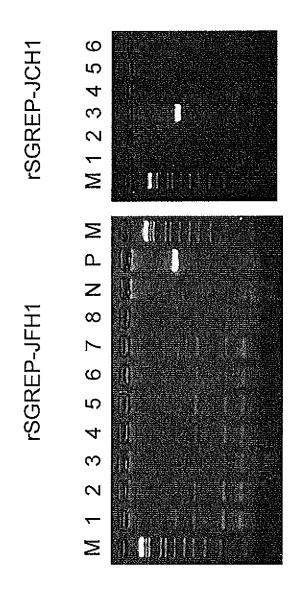
[Figure 7]



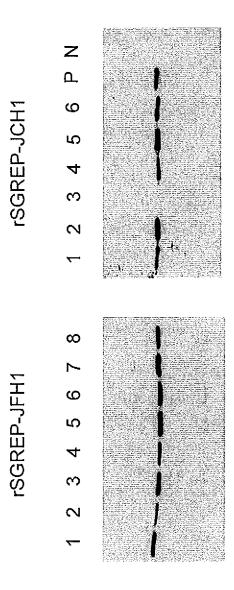
[Figure 8]



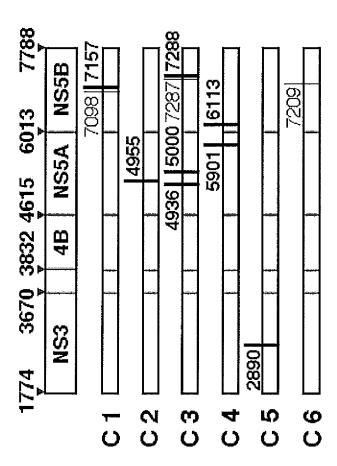
[Figure 9]



[Figure 10]



[Figure 11]



[Title of Document] ABSTRACT

[Abstract]

[Technical Problem] An object is to provide a replicon RNA that is derived from HCV of a different genotype from genotype 1b.

[Technical Solution] A replicon RNA comprising a nucleotide sequence at least containing the 5' untranslated region, the sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein, and the 3' untranslated region on the genomic RNA of hepatitis C virus of genotype 2a is provided.

[Selected Drawing] None